

(19)



Europäisches Patentamt

European Patent Office

Office européen des brevets



(11)

EP 0 953 638 A1

(12)

EUROPEAN PATENT APPLICATION

(43) Date of publication:
03.11.1999 Bull. tin 1999/44

(21) Application number: 98400565.2

(22) Date of filing: 11.03.1998

(51) Int. Cl.⁶: **C12N 15/12**, **C07K 14/705**,
C07K 16/28, **A61K 38/17**,
C12Q 1/68, **G01N 33/50**,
A61K 31/70

(84) Designated Contracting States:
**AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC
NL PT SE**

Designated Extension States:
AL LT LV MK RO SI

(71) Applicant: **SYNTHELABO**
92350 Le Plessis Robinson (FR)

(72) Inventors:
• **Partiseti, Michel**
92160 Antony (FR)

• **Renard, Stéphane**
92500 Rueil Malmaison (FR)

(74) Representative:
Thouret-Lemaître, Elisabeth et al
Sanofi-Synthelabo
Service Brevets
174, avenue de France
75013 Paris (FR)

(54) **Human vanilloid receptor-like cation channel**

(57) hVRCC polypeptides and polynucleotides and methods for producing such polypeptides by recombinant techniques are disclosed. Also disclosed are methods for utilizing hVRCC polypeptides and polynucleotides in the design of protocols for the treatment of acute and chronic inflammation, acute and chronic pain, brain diseases, abnormal proliferation and cancer, ulcer, autoimmune diseases, to mimic or antagonize effect of endogenous neurotransmitters and hormones, to inhibit graft rejection by promoting immunosuppression, among others and diagnostic assays for such conditions.

EP 0 953 638 A1

Description**FIELD OF INVENTION**

5 [0001] This invention relates to newly identified polynucleotides, polypeptides encoded by such polynucleotides, to the use of such polynucleotides and polypeptides and to their production. More particularly, the polynucleotides and polypeptides of the present invention relate to vanilloid receptor-like channel family, hereinafter referred to as hVRCC (Human Vanilloid Receptor-like Cation Channel). The invention also relates to inhibiting or activating the action of such polynucleotides and polypeptides.

BACKGROUND OF THE INVENTION

10 [0002] Vanilloids are natural compounds which are known to trigger cation permeability in subpopulations of peripheral neurons. These ones, also called "nociceptors", are involved in physiological processes such as transmission to the central nervous system of noxious stimuli, said stimuli being mechanical, chemical or thermal (Jessel and Kelly, 1991, pp 385-399, Principal of Neural Sciences, third edition, edited by Kandel et al.). Recently, a new cation channel was discovered and isolated from rat (Caterina et al., 1997, Nature 389 pp 816-824). This channel is activated by vanilloids such as capsaicin and resiniferatoxin and is highly expressed in adult dorsal root ganglia. This channel has also been shown to have significant structural similarities with the "store-operated" calcium channel family i. e., six putative transmembrane domains. A major functional characteristic of this capsaicin-gated conductance is that it is highly selective for the divalent cation calcium even if it is also permeant to magnesium and monovalent cations such as sodium, potassium and cesium. This indicates that these channels have an interesting potential as therapeutic targets. Clearly there is a need for identification and characterization of further channels which can play a role in (i) preventing, ameliorating or correcting dysfunctions or diseases, including, but not limited to, acute and chronic inflammation, acute and chronic pain, brain diseases, ulcer, abnormal proliferation and cancer, autoimmune diseases and (ii) mimicking or antagonizing effect of endogenous neurotransmitters and hormones and inhibiting graft rejection by promoting immunosuppression.

SUMMARY OF THE INVENTION

30 [0003] In one aspect, the invention relates to hVRCC polypeptides and recombinant materials and methods for their production. Another aspect of the invention relates to methods for using such hVRCC polypeptides and polynucleotides. Such uses include the treatment of acute and chronic inflammation, acute and chronic pain, brain diseases, abnormal proliferation and cancer, ulcer, to mimic or antagonize effect of endogenous neurotransmitters and hormones, or to diagnose or treat any disorder related to abnormal expression of these hVRCC polypeptides, among others. In still another aspect, the invention relates to methods to identify agonists and antagonists using the materials provided by the invention, and treating conditions associated with hVRCC imbalance with the identified compounds. Yet another aspect of the invention relates to diagnostic assays for detecting diseases associated with inappropriate hVRCC activity or levels.

DESCRIPTION OF THE INVENTION**Definitions**

45 [0004] The following definitions are provided to facilitate understanding of certain terms used frequently herein-below.

[0005] "hVRCC" refers, among others, to a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:2, or an allelic variant thereof.

50 [0006] "Receptor Activity" or "Channel Activity" or "Biological Activity of the Receptor" or "Biological Activity of the Channel" refers to the metabolic or physiologic function of said hVRCC including similar activities or improved activities or these activities with decreased undesirable side-effects. Also included are antigenic and immunogenic activities of said hVRCC.

[0007] "hVRCC gene" refers to a polynucleotide comprising the nucleotide sequence set forth in SEQ ID NO:1 or allelic variants thereof and/or their complements.

55 [0008] "Antibodies" as used herein includes polyclonal and monoclonal antibodies, chimeric, single chain, and humanized antibodies, as well as Fab fragments, including the products of a Fab or other immunoglobulin expression library.

[0009] "Isolated" means altered "by the hand of man" from the natural state. If an "isolated" composition or substance occurs in nature, it has been changed or removed from its original environment, or both. For example, a polynucleotide

or a polypeptide naturally present in a living animal is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein.

[0010] "Polynucleotide" generally refers to any polyribonucleotide or polydeoxyribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. "Polynucleotides" include, without limitation single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, "polynucleotide" refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The term polynucleotide also includes DNAs or RNAs containing one or more modified bases and DNAs or RNAs with backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications has been made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically or metabolically modified forms of polynucleotides as typically found in nature, as well as the chemical forms of DNA and RNA characteristic of viruses and cells. "Polynucleotide" also embraces relatively short polynucleotides, often referred to as oligonucleotides.

[0011] "Polypeptide" refers to any peptide or protein comprising two or more amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres. "Polypeptide" refers to both short chains, commonly referred to as peptides, oligopeptides or oligomers, and to longer chains, generally referred to as proteins. Polypeptides may contain amino acids other than the 20 gene-encoded amino acids. "Polypeptides" include amino acid sequences modified either by natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched and branched cyclic polypeptides may result from post-translational natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cystine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York, 1993 and Wold, F., Posttranslational Protein Modifications: Perspectives and Prospects, pgs. 1-12 in POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, 1983; Seifter *et al.*, "Analysis for protein modifications and nonprotein cofactors", *Meth Enzymol* (1990) 182:626-646 and Rattan *et al.*, "Protein Synthesis: Posttranslational Modifications and Aging", *Ann NY Acad Sci* (1992) 663:48-62.

[0012] "Variant" as the term is used herein, is a polynucleotide or polypeptide that differs from a reference polynucleotide or polypeptide respectively, but retains essential properties. A typical variant of a polynucleotide differs in nucleotide sequence from another, reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. A typical variant of a polypeptide differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. A variant of a polynucleotide or polypeptide may be naturally occurring such as an allelic variant, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of polynucleotides and polypeptides may be made by mutagenesis techniques or by direct synthesis.

[0013] "Identity" is a measure of the identity of nucleotide sequences or amino acid sequences. In general, the sequences are aligned so that the highest order match is obtained. "Identity" *per se* has an art-recognized meaning and can be calculated using published techniques. See, e.g.: (COMPUTATIONAL MOLECULAR BIOLOGY, Lesk, A.M., ed., Oxford University Press, New York, 1988; BIOCOMPUTING: INFORMATICS AND GENOME PROJECTS, Smith, D.W., ed., Academic Press, New York, 1993; COMPUTER ANALYSIS OF SEQUENCE DATA, PART I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; SEQUENCE ANALYSIS IN MOLECULAR BIOLOGY, von Heinje, G., Academic Press, 1987; and SEQUENCE ANALYSIS PRIMER, Gribskov, M. and Devereux, J., eds., M Stockton Press,

New York, 1991). While there exist a number of methods to measure identity between two polynucleotide or polypeptide sequences, the term "identity" is well known to skilled artisans (Carillo, H., and Lipton, D., *SIAM J Applied Math* (1988) 48:1073). Methods commonly employed to determine identity or similarity between two sequences include, but are not limited to, those disclosed in Guide to Huge Computers, Martin J. Bishop, ed., Academic Press, San Diego, 1994, and Carillo, H., and Lipton, D., *SIAM J Applied Math* (1988) 48:1073. Methods to determine identity and similarity are codified in computer programs. Preferred computer program methods to determine identity and similarity between two sequences include, but are not limited to, GCG program package (Devereux, J., et al., *Nucleic Acids Research* (1984) 12(1):387), BLASTP, BLASTN, FASTA (Atschul, S.F. et al., *J Molec Biol* (1990) 215:403).

[0014] As an illustration, by a polynucleotide having a nucleotide sequence having at least, for example, 95% "identity" to a reference nucleotide sequence of SEQ ID NO: 1 is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence of SEQ ID NO: 1. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. These mutations of the reference sequence may occur at the 5 or 3 terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence.

[0015] Similarly, by a polypeptide having an amino acid sequence having at least, for example, 95% "identity" to a reference amino acid sequence of SEQ ID NO:2 is intended that the amino acid sequence of the polypeptide is identical to the reference sequence except that the polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the reference amino acid of SEQ ID NO: 2. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a reference amino acid sequence, up to 5% of the amino acid residues in the reference sequence may be deleted or substituted with another amino acid, or a number of amino acids up to 5% of the total amino acid residues in the reference sequence may be inserted into the reference sequence. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

30 Polypeptides of the Invention

[0016] In one aspect, the present invention relates to hVRCC polypeptides. The hVRCC polypeptides include the polypeptide of SEQ ID NO:2, as well as polypeptides comprising the amino acid sequence of SEQ ID NO:2, and polypeptides comprising the amino acid sequence which have at least 80% identity to that of SEQ ID NO:2 over its entire length, and still more preferably at least 90% identity, and even still more preferably at least 95% identity to SEQ ID NO: 2. Furthermore, those with at least 97-99% are highly preferred. Also included within hVRCC polypeptides are polypeptides having the amino acid sequence which have at least 80% identity to the polypeptide having the amino acid sequence of SEQ ID NO: 2 over its entire length, and still more preferably at least 90% identity, and even still more preferably at least 95% identity to SEQ ID NO: 2. Furthermore, those with at least 97-99% are highly preferred. Preferably hVRCC polypeptides exhibit at least one biological activity of the receptor.

[0017] The hVRCC polypeptides may be in the form of the "mature" protein or may be a part of a larger protein such as a fusion protein. It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, pro-sequences, sequences which aid in purification such as multiple histidine residues, or an additional sequence for stability during recombinant production.

[0018] Fragments of the hVRCC polypeptides are also included in the invention. A fragment is a polypeptide having an amino acid sequence that entirely is the same as part, but not all, of the amino acid sequence of the aforementioned hVRCC polypeptides. As with hVRCC polypeptides, fragments may be "free-standing," or comprised within a larger polypeptide of which they form a part or region, most preferably as a single continuous region. Representative examples of polypeptide fragments of the invention, include, for example, fragments from about amino acid number 1-20, 21-40, 41-60, 61-80, 81-100, and 101 to the end of hVRCC polypeptide. In this context "about" includes the particularly recited ranges larger or smaller by several, 5, 4, 3, 2 or 1 amino acid at either extreme or at both extremes.

[0019] Preferred fragments include, for example, truncation polypeptides having the amino acid sequence of hVRCC polypeptides, except for deletion of a continuous series of residues that includes the amino terminus, or a continuous series of residues that includes the carboxyl terminus or deletion of two continuous series of residues, one including the amino terminus and one including the carboxyl terminus. Also preferred are fragments characterized by structural or functional attributes such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate bind-

ing region, and high antigenic index regions. Other preferred fragments are biologically active fragments. Biologically active fragments are those that mediate receptor or channel activity, including those with a similar activity or an improved activity, or with a decreased undesirable activity. Also included are those that are antigenic or immunogenic in an animal, especially in a human.

- 5 **[0020]** Preferably, all of these polypeptide fragments retain the biological activity of the receptor, including antigenic activity. Variants of the defined sequence and fragments also form part of the present invention. Preferred variants are those that vary from the referents by conservative amino acid substitutions i.e., those that substitute a residue with another of like characteristics. Typical such substitutions are among Ala, Val, Leu and Ile; among Ser and Thr; among the acidic residues Asp and Glu; among Asn and Gln; and among the basic residues Lys and Arg; or aromatic residues
- 10 Phe and Tyr. Particularly preferred are variants in which several, 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination.

- [0021]** The hVRCC polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well under-
- 15 stood in the art.

Polynucleotides of the Invention

- [0022]** Another aspect of the invention relates to hVRCC polynucleotides. hVRCC polynucleotides include isolated
- 20 polynucleotides which encode the hVRCC polypeptides and fragments, and polynucleotides closely related thereto. More specifically, hVRCC polynucleotides of the invention include a polynucleotide comprising the nucleotide sequence set forth in SEQ ID NO:1 encoding a hVRCC polypeptide of SEQ ID NO: 2, and polynucleotide having the particular sequence of SEQ ID NO:1. hVRCC polynucleotides further include a polynucleotide comprising a nucleotide sequence that has at least 80% identity to a nucleotide sequence encoding the hVRCC polypeptide of SEQ ID NO:2 over its entire
- 25 length, and a polynucleotide that is at least 80% identical to that having SEQ ID NO:1 over its entire length. In this regard, polynucleotides at least 90% identical are particularly preferred, and those with at least 95% are especially preferred. Furthermore, those with at least 97% are highly preferred and those with at least 98-99% are most highly preferred, with at least 99% being the most preferred. Also included under hVRCC polynucleotides are a nucleotide sequence which has sufficient identity to a nucleotide sequence contained in SEQ ID NO:1 or contained in the cDNA
- 30 insert in the plasmid deposited with the ATCC Deposit number 209625 to hybridize under conditions useable for amplification or for use as a probe or marker. Moreover, hVRCC polynucleotide includes nucleotide sequences having at least 80% identity to a nucleotide sequence encoding the hVRCC polypeptide expressed by the cDNA insert deposited at the ATCC with Deposit Number 209625, and a nucleotide sequence comprising at least 15 contiguous nucleotides of such cDNA insert. In this regard, polynucleotides at least 90% identical are particularly preferred, and those with at
- 35 least 95% are especially preferred. Furthermore, those with at least 97% are highly preferred and those with at least 98-99% are most highly preferred, with at least 99% being the most preferred. The invention also provides polynucleotides which are complementary to all the above hVRCC polynucleotides.

- [0023]** A deposit containing a human hVRCC cDNA has been deposited with the American Type Culture Collection (ATCC), 12301 Park Lawn Drive, Rockville, Maryland 20852, USA, on February 10, 1998, and assigned ATCC Deposit
- 40 Number 209625. The deposited material (clone) is a DH5-a strain containing the expression vector bluescript (pBS-SK, Stratagene) that further contains the hVRCC cDNA, referred to as "hVRCC" upon deposit. The cDNA insert is within EcoRI-XhoI site(s) in the vector. The nucleotide sequence of the polynucleotides contained in the deposited material, as well as the amino acid sequence of the polypeptide encoded thereby, are controlling in the event of any conflict with any description of sequences herein. The deposit has been made under the terms of the Budapest Treaty on the inter-
- 45 national recognition of the deposit of micro-organisms for purposes of patent procedure. The strain will be irrevocably and without restriction or condition released to the public upon the issuance of patent.

- [0024]** hVRCC of the invention is structurally related to other proteins of the store-operated calcium channel family, as shown by the results of sequencing the cDNA of Table 1 (SEQ ID NO:1) encoding human hVRCC. The cDNA sequence of SEQ ID NO:1 contains an open reading frame (nucleotide number 402 to 2696) encoding a polypeptide of
- 50 763 amino acids of SEQ ID NO:2. Amino acid sequence of Table 2 (SEQ ID NO:2) has about 49.6% identity (using blastP version 1.4, GCG program package) in 681 amino acid residues with the rat vanilloid receptor subtype 1 (VR1) protein (accession number: AF029310, Caterina et al., 1997, Nature 389, pp816-824). Nucleotide sequence of Table 1 (SEQ ID NO:1) has about 61% identity (using blastN version 1.4, GCG program package) in 2069 (166 to 2235) nucleotide residues with the VR1 cDNA (accession number: AF029310).

55

Table 1*

5

10

15

20

25

30

35

40

45

```

GGCTAGCCTGTCTGACAGGGGAGAGTTAAGCTCCCGTTCTCCACCGTGCCGGCTGGCCAGGT
GGGCTGAGGGTGACCGAGAGACCAGAACCTGCTTGCTGGAGCTTAGTGCTCAGAGCTGGGGAG
GGAGGTTCCGCCGCTCCTCTGCTGTACGCGCCGGCAGCCCTCCCGGCTTCACTTCCTCCCGC
AGCCCCCTGCTACTGAGAAGCTCCGGGATCCAGCAGCCGCCACGCCCTGGCCTCAGCCTGCGG
GGCTCCAGTCAGGCCAACCCGACGCGCAGCTGGGAGGAAGACAGGACCCTTGACATCTCCAT
CTGCACAGAGGTCTGGCTGGACCGAGCAGCCTCCTCCTCTAGGATGACCTCACCTCCAGC
TCTCCAGTTTTCAGGTTGGAGACATTAGATGGAGGCCAAGAAGATGGCTCTGAGGCGGACAGA
GGAAAGCTGGATTTTGGGAGCGGGCTGCCTCCCATGGAGTCACAGTTCCAGGGCGAGGACCGG
AAATTTCGCCCCCTCAGATAAGAGTCAACCTCAACTACCGAAAGGGAACAGGTGCCAGTCAGCCG
GATCCAAACCGATTTCACCGAGATCGGCTCTTCAATGCGGTCTCCCGGGGTGTCCCGAGGAT
CTGGCTGGACTTCCAGAGTACCTGAGCAAGACCAGCAAGTACCTCACCGACTCGGAATACACA
GAGGGCTCCACAGGTAAGACGTGCCTGATGAAGGCTGTGCTGAACCTTAAGGACGGAGTCAAT
GCCTGCATTCTGCCACTGCTGCAGATCGACAGGGACTCTGGCAATCCTCAGCCCCCTGGTAAAT
GCCAGTGACAGATGACTATTACCGAGGCCACAGCGCTCTGCACATCGCCATTGAGAAGAGG
AGTCTGCAGTGTGTGAAGCTCCTGGTGGAGAATGGGGCCAATGTGCATGCCCGGGCTGCGGC
CGCTTCTTCCAGAAGGGCCAAGGGACTTGCTTTTATTTTCGGTGAGCTACCCCTCTCTTTGGCC
GCTTGACCAAGCAGTGGGATGTGGAAGCTACCTCCTGGAGAACCACACCAGCCCGCCAGC
CTGCAGGCCACTGACTCCCAGGGCAACACAGTCTGCATGCCCTAGTGATGATCTCGGACAAC
TCAGCTGAGAACATTGCACTGGTGACCAGCATGTATGATGGGCTCCTCCAAGCTGGGGCCCGC
CTCTGCCCTACCGTGCAGCTTGAGGACATCCGCAACCTGCAGGATCTCACGCCCTCTGAAGCTG
GCCGCCAAGGAGGGCAAGATCGAGATTTTCAGGCACATCTGCAGCGGGAGTTTTCAGGACTG
AGCCACCTTTCCGAAAGTTTACCAGTGGTGCTATGGGCCTGTCCGGGTGTGCTGTATGAC
CTGGCTTCTGTGGACAGCTGTGAGGAGAACTCAGTGCTGGAGATCATTGCCTTTCATTGCAAG
AGCCCGCACCGACACCGAATGGTCTGTTTTGGAGCCCTGAACAACTGCTGCAGGCGAAATGG
GATCTGCTCATCCCCAAGTTCTTCTTAAACTTCTGTGTAATCTGATCTACATGTTTCACTTC
ACCGCTGTTGCCATACCATCAGCCTACCTGAAGAAGGCCGCCCTCACCTGAAAGCGGAGGT
GGAAACTCCATGCTGCTGACGGGCCACATCCTTATCCTGCTAGGGGGGATCTACCTCCTCGTG
GGCCAGCTGTGGTACTTCTGGCGGCGCCACGTGTTTCTGATCTCGTTTATAGACAGCTAC
TTTGAAATCCTTCTTCTGTTCCAGGCCCTGCTCACAGTGGTGTCCAGGTGCTGTGTTTCTG
GCCATCGAGTGGTACCTGCCCTGCTTGTGCTGCGCTGGTGGCTGGCTGAACCTGCTT
TACTATACAGTGGCTTCCAGCACACAGGCATCTACAGTGTGATGATCCAGAAGGTCATCCTG
CGGGACCTGCTGCGCTTCTTCTGATCTACTTAGTCTTCTCTTTTCGGCTTCGCTGTAGCCCTG
GTGAGCCTGAGCCAGGAGGCTTGGCGCCCCGAAGCTCCTACAGGCCCAATGCCACAGAGTCA
GTGCAGCCCATGGAGGGACAGGAGGACGAGGGCAACGGGGCCAGTACAGGGGTATCCTGGAA
GCCTCCTTGGAGCTCTTCAAATTCACCATCGGCATGGGCGAGCTGGCCTTCCAGGAGCAGCTG
CACTTCCGCGGCATGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG

```

* A nucleotide sequence of a human hVRCC. SEQ ID NO: 1.

50 [0025] Allelic variants of this sequence have been identified such as a t in position 374, a g in position 750, a c in position 787, and an agg insertion after position 1612, resulting in a glutamine amino-acid insertion in the corresponding position of the protein.

55

Table 2^b

5 MTSPSSSPVFRLETL DGGQEDGSEADRGKLD FGSGLP PMESQFQGEDRK FAPQIRVN
 LNYRKGTGASQ PDPNRFDRDLF NAVSRGVPEDLAGL PEYLSKTSKY L TDSEYTEGS
 10 TGKTCLMKAVLNLK DGVNACILPLLQID RDSGNPQPLVNAQCTDDYYRGHSALHIAI
 EKRS LQC VKLLVENGANV HARACGRFFQKGQGT CFYFGELPLSLAACTKQWDVVSYL
 LENPHQPASLQATDSQGNTVLHALVMISDNSAENIALVTSMYDGLLQAGARLCPTVQ
 LEDIRNLQDLT PLKLAKEGKIEIFRHILQREFSGLSHLSRKFT EW CYGPVRVSLYD
 15 LASVDSCEENS VLEIIAFHCKSPHRHRMVVLEPLNKLLQAKWDL LIPKFFLNFLCNL
 IYMFIFTAVAYHQPTL KKAAPHLKAEVGN SMLLTGHILILLGGIYLLVGQLWYFWRR
 H VFIWISFIDSYFEILFLFQALLTVVSQVLCFLAIEWYLPLLVSALVLGWLNL LYYT
 RGFQHTGIYSVMIQKVILRDLLRFLLIYLVFLFGFAVALVSLSQEAWRPEAPTGPNA
 20 TESVQPMEGQEDEGN GAQYRGILEASLELFKFTIGMGELAFQEQLHFRGMVLLLLLLA
 YVLLTYILLNMLIALMSETVNSVATDSWSIWKLQKAISVLEMENGYWWCRKKQRAG
 VMLTVGTKPDGSPDERWC FRVEEVN WASWEQTLPTLCEDPSGAGVPRTLENPVLASP
 25 PKEDEDGASEENYVPVQLLSN

^b An amino acid sequence of a human hVRCC. SEQ ID NO: 2.

30 [0026] One polynucleotide of the present invention encoding hVRCC may be obtained using standard cloning and screening, from a cDNA library derived from mRNA in macrophages using the expressed sequence tag (EST) analysis (Adams, M.D., *et al. Science* (1991) 252:1651-1656; Adams, M.D. *et al., Nature*, (1992) 355:632-634; Adams, M.D., *et al., Nature* (1995) 377 Supp:3-174). Polynucleotides of the invention can also be obtained from natural sources such as genomic DNA libraries or can be synthesized using well known and commercially available techniques.

35 [0027] The nucleotide sequence encoding hVRCC polypeptide of SEQ ID NO:2 may be identical to the polypeptide encoding sequence contained in Table 1 (nucleotide number 361 to 2649 of SEQ ID NO:1), or it may be a sequence, which as a result of the redundancy (degeneracy) of the genetic code, also encodes the polypeptide of SEQ ID NO:2.

[0028] When the polynucleotides of the invention are used for the recombinant production of hVRCC polypeptide, the polynucleotide may include the coding sequence for the mature polypeptide or a fragment thereof, by itself; the coding sequence for the mature polypeptide or fragment in reading frame with other coding sequences, such as those encoding a leader or secretory sequence, a pre-, or pro- or prepro- protein sequence, or other fusion peptide portions. For example, a marker sequence which facilitates purification of the fused polypeptide can be encoded. In certain preferred embodiments of this aspect of the invention, the marker sequence is a hexa-histidine peptide, as provided in the pQE vector (Qiagen, Inc.) and described in Gentz *et al., Proc Natl Acad Sci USA* (1989) 86:821-824, or is an HA tag. The polynucleotide may also contain non-coding 5' and 3' sequences, such as transcribed, non-translated sequences, splicing and polyadenylation signals, ribosome binding sites and sequences that stabilize mRNA.

45 [0029] Further preferred embodiments are polynucleotides encoding hVRCC variants comprising the amino acid sequence of hVRCC polypeptide of Table 2 (SEQ ID NO:2) in which several, 5-10, 1-5, 1-3, 1-2 or 1 amino acid residues are substituted, deleted or added, in any combination.

50 [0030] The present invention further relates to polynucleotides that hybridize to the herein above-described sequences. In this regard, the present invention especially relates to polynucleotides which hybridize under stringent conditions to the herein above-described polynucleotides. As herein used, the term "stringent conditions" means hybridization will occur only if there is at least 95% and preferably at least 97% identity between the sequences.

[0031] Polynucleotides of the invention, which are identical or sufficiently identical to a nucleotide sequence contained in SEQ ID NO:1 or a fragment thereof, or to the cDNA insert in the plasmid deposited at the ATCC with Deposit Number 209625 or a fragment thereof, may be used as hybridization probes for cDNA and genomic DNA, to isolate full-length cDNAs and genomic clones encoding hVRCC and to isolate cDNA and genomic clones of other genes that have a high sequence similarity to the hVRCC gene. Such hybridisation techniques are known to those of skill in the art. Typically

these nucleotide sequences are 80% identical, preferably 90% identical, more preferably 95% identical to that of the referent. The probes generally will comprise at least 15 nucleotides. Preferably, such probes will have at least 30 nucleotides and may have at least 50 nucleotides. Particularly preferred probes will range between 30 and 50 nucleotides.

[0032] In one embodiment, to obtain a polynucleotide encoding hVRCC polypeptide comprises the steps of screening an appropriate library under stringent hybridization conditions with a labeled probe having the SEQ ID NO: 1 or a fragment thereof; and isolating full-length cDNA and genomic clones containing said polynucleotide sequence. Such hybridization techniques are well known to those of skill in the art. Stringent hybridization conditions are as defined above or alternatively conditions under overnight incubation at 42°C in a solution comprising: 50% formamide, 5xSSC (150mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH7.6), 5x Denhardt's solution, 10 % dextran sulfate, and 20 microgram/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65°C.

[0033] The polynucleotides and polypeptides of the present invention may be employed as research reagents and materials for discovery of treatments and diagnostics to animal and human disease.

15 Vectors, Host Cells, Expression

[0034] The present invention also relates to vectors which comprise a polynucleotide or polynucleotides of the present invention, and host cells which are genetically engineered with vectors of the invention and to the production of polypeptides of the invention by recombinant techniques. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention.

[0035] For recombinant production, host cells can be genetically engineered to incorporate expression systems or portions thereof for polynucleotides of the present invention. Introduction of polynucleotides into host cells can be effected by methods described in many standard laboratory manuals, such as Davis et al., *BASIC METHODS IN MOLECULAR BIOLOGY* (1986) and Sambrook et al., *MOLECULAR CLONING: A LABORATORY MANUAL*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989) such as calcium phosphate transfection, DEAE-dextran mediated transfection, transfection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction or infection.

[0036] Representative examples of appropriate hosts include bacterial cells, such as streptococci, staphylococci, *E. coli*, *Streptomyces* and *Bacillus subtilis* cells; fungal cells, such as yeast cells and *Aspergillus* cells; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, HEK 293 and Bowes melanoma cells; and plant cells.

[0037] A great variety of expression systems can be used. Such systems include, among others, chromosomal, episomal and virus-derived systems, e.g., vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids. The expression systems may contain control regions that regulate as well as engender expression. Generally, any system or vector suitable to maintain, propagate or express polynucleotides to produce a polypeptide in a host may be used. The appropriate nucleotide sequence may be inserted into an expression system by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook et al., *MOLECULAR CLONING, A LABORATORY MANUAL* (*supra*).

[0038] For secretion of the translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signals may be incorporated into the desired polypeptide. These signals may be endogenous to the polypeptide or they may be heterologous signals.

[0039] If the hVRCC polypeptide is to be expressed for use in screening assays, generally, it is preferred that the polypeptide be produced at the surface of the cell. In this event, the cells may be harvested prior to use in the screening assay. If hVRCC polypeptide is secreted into the medium, the medium can be recovered in order to recover and purify the polypeptide; if produced intracellularly, the cells must first be lysed before the polypeptide is recovered.

[0040] hVRCC polypeptides can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography is employed for purification. Well known techniques for refolding proteins may be employed to regenerate active conformation when the polypeptide is denatured during isolation and or purification.

55 Diagnostic Assays

[0041] This invention also relates to the use of hVRCC polynucleotides for use as diagnostic reagents. Detection of

a mutated form of hVRCC gene associated with a dysfunction will provide a diagnostic tool that can add to or define a diagnosis of a disease or susceptibility to a disease which results from under-expression, over-expression or altered expression of hVRCC. Individuals carrying mutations in the hVRCC gene may be detected at the DNA level by a variety of techniques.

5 [0042] Nucleic acids for diagnosis may be obtained from a subject's cells, such as from blood, urine, saliva, tissue biopsy or autopsy material. The genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR or other amplification techniques prior to analysis. RNA or cDNA may also be used in similar fashion. Deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal genotype. Point mutations can be identified by hybridizing amplified DNA to labeled hVRCC nucleotide sequences. Perfectly
10 matched sequences can be distinguished from mismatched duplexes by RNase digestion or by differences in melting temperatures. DNA sequence differences may also be detected by alterations in electrophoretic mobility of DNA fragments in gels, with or without denaturing agents, or by direct DNA sequencing. See, e.g., Myers *et al.*, *Science* (1985) 230:1242. Sequence changes at specific locations may also be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method. See Cotton *et al.*, *Proc Natl Acad Sci USA* (1985) 85: 4397-
15 4401. In another embodiment, an array of oligonucleotides probes comprising hVRCC nucleotide sequence or fragments thereof can be constructed to conduct efficient screening of e.g., genetic mutations. Array technology methods are well known and have general applicability and can be used to address a variety of questions in molecular genetics including gene expression, genetic linkage, and genetic variability. (See for example: M.Chee *et al.*, *Science*, Vol 274, pp 610-613 (1996)).

20 [0043] The diagnostic assays offer a process for diagnosing or determining a susceptibility to inflammation and pain (both acute and chronic), brain diseases, abnormal proliferation and cancer, autoimmune diseases, through detection of mutation in the hVRCC gene by the methods described.

[0044] In addition, inflammation and pain (both acute and chronic), brain diseases, abnormal proliferation and cancer, can be diagnosed by methods comprising determining from a sample derived from a subject an abnormally decreased
25 or increased level of hVRCC polypeptide or hVRCC mRNA. Decreased or increased expression can be measured at the RNA level using any of the methods well known in the art for the quantitation of polynucleotides, such as, for example, PCR, RT-PCR, RNase protection, Northern blotting and other hybridization methods. Assay techniques that can be used to determine levels of a protein, such as a hVRCC, in a sample derived from a host are well-known to those of skill in the art. Such assay methods include radioimmunoassays, competitive-binding assays, Western Blot analysis and
30 ELISA assays.

Chromosome Assays

[0045] The nucleotide sequences of the present invention are also valuable for chromosome identification. The
35 sequence is specifically targeted to and can hybridize with a particular location on an individual human chromosome. The mapping of relevant sequences to chromosomes according to the present invention is an important first step in correlating those sequences with gene associated disease. Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found, for example, in V. McKusick, *Mendelian Inheritance in Man* (available on line through Johns Hopkins University Welch Medical Library). The relationship between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (coinheritance of physically adjacent genes).

[0046] The differences in the cDNA or genomic sequence between affected and unaffected individuals can also be determined. If a mutation is observed in some or all of the affected individuals but not in any normal individuals, then the mutation is likely to be the causative agent of the disease.
45

Antibodies

[0047] The polypeptides of the invention or their fragments or analogs thereof, or cells expressing them can also be used as immunogens to produce antibodies immunospecific for the hVRCC polypeptides. The term "immunospecific"
50 means that the antibodies have substantial greater affinity for the polypeptides of the invention than their affinity for other related polypeptides in the prior art.

[0048] Antibodies generated against the hVRCC polypeptides can be obtained by administering the polypeptides or epitope-bearing fragments, analogs or cells to an animal, preferably a nonhuman, using routine protocols. For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can
55 be used. Examples include the hybridoma technique (Kohler, G. and Milstein, C., *Nature* (1975) 256:495-497), the trioma technique, the human B-cell hybridoma technique (Kozbor *et al.*, *Immunology Today* (1983) 4:72) and the EBV-hybridoma technique (Cole *et al.*, *MONOCLONAL ANTIBODIES AND CANCER THERAPY*, pp. 77-96, Alan R. Liss, Inc., 1985).

[0049] Techniques for the production of single chain antibodies (U.S. Patent No. 4,946,778) can also be adapted to produce single chain antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms including other mammals, may be used to express humanized antibodies.

[0050] The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptide or to purify the polypeptides by affinity chromatography.

[0051] Antibodies against hVRCC polypeptides may also be employed to treat cerebral and cardiac and renal ischemias, brain and cardiac diseases, inflammation, pain, to mimic or antagonize effect of endogenous neurotransmitters and hormones, or to diagnose or treat any disorder related to abnormal expression of said hVRCC polypeptides, among others.

Vaccines and Immunological products

[0052] Another aspect of the invention relates to a method for inducing an immunological response in a mammal which comprises inoculating the mammal with hVRCC polypeptide, or a fragment thereof, adequate to produce antibody and/or T cell immune response to protect said animal from inflammation and pain (both acute and chronic), brain diseases, ulcer, abnormal proliferation and cancer, autoimmune diseases, to mimic or antagonize effect of endogenous neurotransmitters and hormones, or to diagnose or treat any disorder related to abnormal expression of the hVRCC polypeptide, among others. Yet another aspect of the invention relates to a method of inducing immunological response in a mammal which comprises, delivering hVRCC polypeptide via a vector directing expression of hVRCC polynucleotide *in vivo* in order to induce such an immunological response to produce antibody to protect said animal from diseases.

[0053] Further aspect of the invention relates to an immunological/vaccine formulation (composition) which, when introduced into a mammalian host, induces an immunological response in that mammal to a hVRCC polypeptide wherein the composition comprises a hVRCC polypeptide or hVRCC gene. The vaccine formulation may further comprise a suitable carrier. Since hVRCC polypeptide may be broken down in the stomach, it is preferably administered parenterally (including subcutaneous, intramuscular, intravenous, intradermal etc. injection). Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation isotonic with the blood of the recipient; and aqueous and non-aqueous sterile suspensions which may include suspending agents or thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example, sealed ampoules and vials and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use. The vaccine formulation may also include adjuvant systems for enhancing the immunogenicity of the formulation, such as oil-in water systems and other systems known in the art. The dosage will depend on the specific activity of the vaccine and can be readily determined by routine experimentation.

Screening Assays

[0054] The hVRCC polypeptide of the present invention may be employed in a screening process for compounds which bind the channel and which activate (agonists) or inhibit activation of (antagonists) the channel polypeptide of the present invention. Thus, polypeptides of the invention may also be used to assess the binding of small molecule substrates and ligands in, for example, cells, cell-free preparations, chemical libraries, and natural product mixtures. These substrates and ligands may be natural substrates and ligands or may be structural or functional mimetics. See Coligan *et al.*, *Current Protocols in Immunology* 1(2):Chapter 5 (1991).

[0055] hVRCC polypeptides are implicated in many biological functions, and possibly pathologies. Accordingly, it is desirous to find compounds and drugs which stimulate hVRCC on the one hand and which can inhibit the function of hVRCC on the other hand. In general, agonists are employed for therapeutic and prophylactic purposes for such conditions as to mimic effect of endogenous neurotransmitters and hormones. Antagonists may be employed for a variety of therapeutic and prophylactic purposes for such conditions as inflammation and pain (both acute and chronic), brain diseases, ulcer, abnormal proliferation and cancer, autoimmune diseases, to antagonize effect of endogenous neurotransmitters and hormones and to inhibit graft rejection by promoting immunosuppression.

[0056] In general, such screening procedures involve producing appropriate cells which express the receptor polypeptide of the present invention on the surface thereof. Such cells include cells from mammals, yeast, *Drosophila* or *E. coli*. Cells expressing the receptor (or cell membrane containing the expressed receptor) are then contacted with a test compound to observe binding, or stimulation or inhibition of a functional response.

[0057] The assays may simply test binding of a candidate compound wherein adherence to the cells bearing the channel is detected by means of a label directly or indirectly associated with the candidate compound or in an assay involving competition with a labeled competitor. Further, these assays may test whether the candidate compound results in a signal generated by activation of the channel, using detection systems appropriate to the cells bearing the

channel at their surfaces. Inhibitors of activation are generally assayed in the presence of a known agonist and the effect on activation by the agonist by the presence of the candidate compound is observed. Standard methods for conducting such screening assays are well understood in the art.

[0058] The recording of hVRCC channel activity may be carried out either by membrane voltage analysis of transfected cells or microinjected xenopus oocytes, directly (patch-clamp for example) or indirectly (fluorescent probes sensitive to changes of intracellular free calcium concentration such as fura-2 and calcium green, Molecular Probes). The assays may simply test binding of a candidate compound wherein adherence to the cells bearing the channel is detected by means of a label directly or indirectly associated with the candidate compound or in an assay involving competition with a labeled competitor. Further, these assays may test whether the candidate compound results in a signal generated by activation of the channel, using detection systems appropriate to the cells bearing the channel at their surfaces. Inhibitors of activation are generally assayed in the presence of a known agonist and the effect on activation by the agonist by the presence of the candidate compound is observed. Standard methods for conducting such screening assays are well understood in the art.

[0059] Examples of potential hVRCC antagonists include antibodies or, in some cases, oligonucleotides or proteins which are closely related to the ligand of the hVRCC, e.g., a fragment of the ligand, or small molecules which bind to the receptor but do not elicit a response, so that the activity of the channel is prevented.

Prophylactic and Therapeutic Methods

[0060] This invention provides methods of treating abnormal conditions related to both an excess of and insufficient amounts of hVRCC activity.

[0061] If the activity of hVRCC is in excess, several approaches are available. One approach comprises administering to a subject an inhibitor compound (antagonist) as hereinabove described along with a pharmaceutically acceptable carrier in an amount effective to inhibit activation by blocking binding of ligands to the hVRCC, or by inhibiting a second signal, and thereby alleviating the abnormal condition.

[0062] In another approach, soluble forms of hVRCC polypeptides still capable of binding the ligand in competition with endogenous hVRCC may be administered. Typical embodiments of such competitors comprise fragments of the hVRCC polypeptide.

[0063] In still another approach, expression of the gene encoding endogenous hVRCC can be inhibited using expression blocking techniques. Known such techniques involve the use of antisense sequences, either internally generated or separately administered. See, for example, O'Connor, *J Neurochem* (1991) 56:560 in Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988). Alternatively, oligonucleotides which form triple helices with the gene can be supplied. See, for example, Lee *et al.*, *Nucleic Acids Res* (1979) 6:3073; Cooney *et al.*, *Science* (1988) 241:456; Dervan *et al.*, *Science* (1991) 251:1360. These oligomers can be administered *per se* or the relevant oligomers can be expressed *in vivo*.

[0064] For treating abnormal conditions related to an under-expression of hVRCC and its activity, several approaches are also available. One approach comprises administering to a subject a therapeutically effective amount of a compound which activates hVRCC, i.e., an agonist as described above, in combination with a pharmaceutically acceptable carrier, to thereby alleviate the abnormal condition. Alternatively, gene therapy may be employed to effect the endogenous production of hVRCC by the relevant cells in the subject. For example, a polynucleotide of the invention may be engineered for expression in a replication defective retroviral vector, as discussed above. The retroviral expression construct may then be isolated and introduced into a packaging cell transduced with a retroviral plasmid vector containing RNA encoding a polypeptide of the present invention such that the packaging cell now produces infectious viral particles containing the gene of interest. These producer cells may be administered to a subject for engineering cells *in vivo* and expression of the polypeptide *in vivo*. For overview of gene therapy, see Chapter 20, Gene Therapy and other Molecular Genetic-based Therapeutic Approaches, (and references cited therein) in Human Molecular Genetics, T. Strachan and A. P. Read, BIOS Scientific Publishers Ltd (1996).

Formulation and Administration

[0065] Peptides, such as the soluble form of hVRCC polypeptides, and agonists and antagonist peptides or small molecules, may be formulated in combination with a suitable pharmaceutical carrier. Such formulations comprise a therapeutically effective amount of the polypeptide or compound, and a pharmaceutically acceptable carrier or excipient. Such carriers include but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol, and combinations thereof. Formulation should suit the mode of administration, and is well within the skill of the art. The invention further relates to pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients of the aforementioned compositions of the invention.

[0066] Polypeptides and other compounds of the present invention may be employed alone or in conjunction with

other compounds, such as therapeutic compounds.

[0067] Preferred forms of systemic administration of the pharmaceutical compositions include injection, typically by intravenous injection. Other injection routes, such as subcutaneous, intramuscular, or intraperitoneal, can be used. Alternative means for systemic administration include transmucosal and transdermal administration using penetrants such as bile salts or fusidic acids or other detergents. In addition, if properly formulated in enteric or encapsulated formulations, oral administration may also be possible. Administration of these compounds may also be topical and/or localized, in the form of salves, pastes, gels and the like.

[0068] The dosage range required depends on the choice of peptide, the route of administration, the nature of the formulation, the nature of the subject's condition, and the judgment of the attending practitioner. Suitable dosages, however, are in the range of 0.1-100 µg/kg of subject. Wide variations in the needed dosage, however, are to be expected in view of the variety of compounds available and the differing efficiencies of various routes of administration. For example, oral administration would be expected to require higher dosages than administration by intravenous injection. Variations in these dosage levels can be adjusted using standard empirical routines for optimization, as is well understood in the art.

[0069] Polypeptides used in treatment can also be generated endogenously in the subject, in treatment modalities often referred to as "gene therapy" as described above. Thus, for example, cells from a subject may be engineered with a polynucleotide, such as a DNA or RNA, to encode a polypeptide ex vivo, and for example, by the use of a retroviral plasmid vector. The cells are then introduced into the subject.

Examples

[0070] The examples below are carried out using standard techniques, which are well known and routine to those of skill in the art, except where otherwise described in detail. The examples illustrate, but do not limit the invention.

Example 1

Cloning the Human hVRCC cation Channel

[0071] The sequence of the hVRCC cation channel was first identified by searching a database containing approximately 2 million human ESTs, which was generated using high throughput automated DNA sequence analysis of randomly selected human cDNA clones (Adams, M.D. *et al.*, *Nature* 377:3-174 (1995); Adams, M.D. *et al.*, *Nature* 355:632-634 (1992); and Adams, M.D. *et al.*, *Science* 252:1651-1656 (1991)). Sequence homology comparisons of each EST were performed against the GenBank database using the blastn and tblastn algorithms (Altschul, S.F. *et al.*, *J. Mol. Biol.* 215:403-410 (1990)). A specific homology search using the known rat VR1 amino acid sequence against this human EST database revealed one EST, from a macrophage cDNA library, with approximately 60% similarity to VR1. The sequence comparison suggested that it contained the complete open reading frame of a new protein. Sequence of the gene was confirmed by double strand DNA sequencing using the TaqFs (Perkin Elmer) and the gene was shown to be completely new by a blast search against Genbank release 103. The entire hVRCC coding region containing the EcoRI-XhoI fragments was inserted into the expression vector bluescript (Stratagene).

Example 2

Cloning and Expression of hVRCC in Mammalian Cells

[0072] A typical mammalian expression vector contains the promoter element, which mediates the initiation of transcription of mRNA, the protein coding sequence, and signals required for the termination of transcription and polyadenylation of the transcript. Additional elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription can be achieved with the early and late promoters from SV40, the long terminal repeats (LTRS) from Retroviruses, e.g., RSV, HTLV1, HIV1 and the early promoter of the cytomegalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter). Suitable expression vectors for use in practicing the present invention include, for example, vectors such as PSVL and PMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146) and pBC12MI (ATCC 67109) and pcDNA3 (Invitrogen). Mammalian host cells that could be used include, human HEK 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and mouse L cells and Chinese hamster ovary (CHO) cells.

[0073] Alternatively, the gene can be expressed in stable cell lines that contain the gene integrated into a chromosome. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, zeocin or hygromycin allows the identification and isolation of the transfected cells.

[0074] The transfected gene can also be amplified to express large amounts of the encoded protein. The DHFR (dihy-

drofolate reductase) marker is useful to develop cell lines that carry several hundred or even several thousand copies of the gene of interest. Another useful selection marker is the enzyme glutamine synthase (GS) (Murphy *et al.*, *Biochem. J.* 227:277-279 (1991); Bebbington *et al.*, *Bio/Technology* 10:169-175 (1992)). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a chromosome. Chinese hamster ovary (CHO) and NSO cells are often used for the production of proteins.

[0075] The expression vector pCMVSPORT3.0 contains the strong promoter (CMV) of the Cytomegalovirus. Multiple cloning sites, e.g., with the restriction enzyme cleavage sites EcoRI, XhoI, facilitate the cloning of the gene of interest.

10 Example 3

Tissue distribution of hVRCC mRNA expression

[0076] Northern blot analysis can be carried out to examine hVRCC gene expression in human tissues, using methods described by, among others, Sambrook *et al.*, cited above. A cDNA probe containing the entire nucleotide sequence of the hVRCC protein can be labeled with ³²P using the Rediprime™ DNA labeling system (Amersham Life Science, Arlington, IL), according to manufacturer's instructions. After labeling, the probe can be purified using a CHROMA SPIN- 100™ column (Clontech Laboratories, Inc.), according to manufacturer's protocol number PT1200-1. The purified labeled probe was then used to examine various human tissues for hVRCC mRNA.

[0077] Multiple Tissue Northern (MTN) blots containing various human tissues can be obtained from Clontech and examined with the labeled probe using ExpressHyb™ hybridization solution (Clontech) according to manufacturer's protocol number PT1190-1. Following hybridization and washing, the blots can be mounted and exposed to film at -70°C overnight, and films developed according to standard procedures.

[0078] It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples.

[0079] Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

[0080] The entire disclosure of all publications (including patents, patent applications, journal articles, laboratory manuals, books, or other documents) cited herein are incorporated by reference.

SEQUENCE LISTING

5 (1) GENERAL INFORMATION:

(i) APPLICANT:

(A) NAME: SYNTHELABO

(B) STREET: 22 avenue Galilee

(C) CITY: LE PLESSIS-ROBINSON

10 (E) COUNTRY: FRANCE

(F) POSTAL CODE (ZIP): 92350

(G) TELEPHONE: (33) 1 45 37 56 76

(ii) TITLE OF INVENTION: Human vanilloid receptor-like cation
channel

15 (iii) NUMBER OF SEQUENCES: 2

(iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk

(B) COMPUTER: IBM PC compatible

20 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

25 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2783 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

30 (ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

35 (vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo sapiens

(ix) FEATURE:

(A) NAME/KEY: 5'UTR

40 (B) LOCATION:1..360

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION:361..2649

45 (ix) FEATURE:

(A) NAME/KEY: 3'UTR

(B) LOCATION:2650..2783

(ix) FEATURE:

(A) NAME/KEY: allele

50 (B) LOCATION:replace(374, "t")

(ix) FEATURE:

(A) NAME/KEY: allele

55

(B) LOCATION:replace(750, "g")

(ix) FEATURE:

(A) NAME/KEY: allele

(B) LOCATION:replace(787, "c")

(ix) FEATURE:

(A) NAME/KEY: allele

(B) LOCATION:replace(1612, "cagg")

(D) OTHER INFORMATION:/label= GLUTAMINE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GGCTAGCCTG TCCTGACAGG GGAGAGTTAA GCTCCCGTTC TCCACCGTGC CGGCTGGCCA
60

GGTGGGCTGA GGGTGACCGA GAGACCAGAA CCTGCTTGCT GGAGCTTAGT GCTCAGAGCT
120

GGGGAGGGAG GTTCCGCCGC TCCTCTGCTG TCAGCGCCGG CAGCCCCTCC CGGCTTCACT
180

TCCTCCCGCA GCCCCTGCTA CTGAGAAGCT CCGGGATCCC AGCAGCCGCC ACGCCCTGGC
240

CTCAGCCTGC GGGGCTCCAG TCAGGCCAAC ACCGACGCGC AGCTGGGAGG AAGACAGGAC
300

CCTTGACATC TCCATCTGCA CAGAGGTCCT GGCTGGACCG AGCAGCCTCC TCCTCCTAGG
360

ATG ACC TCA CCC TCC AGC TCT CCA GTT TTC AGG TTG GAG ACA TTA GAT
408

Met Thr Ser Pro Ser Ser Ser Pro Val Phe Arg Leu Glu Thr Leu Asp
1 5 10 15

GGA GGC CAA GAA GAT GGC TCT GAG GCG GAC AGA GGA AAG CTG GAT TTT
456

Gly Gly Gln Glu Asp Gly Ser Glu Ala Asp Arg Gly Lys Leu Asp Phe
20 25 30

GGG AGC GGG CTG CCT CCC ATG GAG TCA CAG TTC CAG GGC GAG GAC CGG
504

Gly Ser Gly Leu Pro Pro Met Glu Ser Gln Phe Gln Gly Glu Asp Arg
35 40 45

AAA TTC GCC CCT CAG ATA AGA GTC AAC CTC AAC TAC CGA AAG GGA ACA
552

Lys Phe Ala Pro Gln Ile Arg Val Asn Leu Asn Tyr Arg Lys Gly Thr
50 55 60

GGT GCC AGT CAG CCG GAT CCA AAC CGA TTT GAC CGA GAT CGG CTC TTC
600

Gly Ala Ser Gln Pro Asp Pro Asn Arg Phe Asp Arg Asp Arg Leu Phe
65 70 75 80

AAT GCG GTC TCC CGG GGT GTC CCC GAG GAT CTG GCT GGA CTT CCA GAG
648

Asn Ala Val Ser Arg Gly Val Pro Glu Asp Leu Ala Gly Leu Pro Glu
 85 90 95
 5 TAC CTG AGC AAG ACC AGC AAG TAC CTC ACC GAC TCG GAA TAC ACA GAG
 696
 Tyr Leu Ser Lys Thr Ser Lys Tyr Leu Thr Asp Ser Glu Tyr Thr Glu
 100 105 110
 10 GGC TCC ACA GGT AAG ACG TGC CTG ATG AAG GCT GTG CTG AAC CTT AAG
 744
 Gly Ser Thr Gly Lys Thr Cys Leu Met Lys Ala Val Leu Asn Leu Lys
 115 120 125
 GAC GGA GTC AAT GCC TGC ATT CTG CCA CTG CTG CAG ATC GAC AGG GAC
 792
 15 Asp Gly Val Asn Ala Cys Ile Leu Pro Leu Leu Gln Ile Asp Arg Asp
 130 135 140
 TCT GGC AAT CCT CAG CCC CTG GTA AAT GCC CAG TGC ACA GAT GAC TAT
 840
 20 Ser Gly Asn Pro Gln Pro Leu Val Asn Ala Gln Cys Thr Asp Asp Tyr
 145 150 155 160
 TAC CGA GGC CAC AGC GCT CTG CAC ATC GCC ATT GAG AAG AGG AGT CTG
 888
 Tyr Arg Gly His Ser Ala Leu His Ile Ala Ile Glu Lys Arg Ser Leu
 165 170 175
 25 CAG TGT GTG AAG CTC CTG GTG GAG AAT GGG GCC AAT GTG CAT GCC CGG
 936
 Gln Cys Val Lys Leu Leu Val Glu Asn Gly Ala Asn Val His Ala Arg
 180 185 190
 30 GCC TGC GGC CGC TTC TTC CAG AAG GGC CAA GGG ACT TGC TTT TAT TTC
 984
 Ala Cys Gly Arg Phe Phe Gln Lys Gly Gln Gly Thr Cys Phe Tyr Phe
 195 200 205
 GGT GAG CTA CCC CTC TCT TTG GCC GCT TGC ACC AAG CAG TGG GAT GTG
 1032
 35 Gly Glu Leu Pro Leu Ser Leu Ala Ala Cys Thr Lys Gln Trp Asp Val
 210 215 220
 GTA AGC TAC CTC CTG GAG AAC CCA CAC CAG CCC GCC AGC CTG CAG GCC
 1080
 40 Val Ser Tyr Leu Leu Glu Asn Pro His Gln Pro Ala Ser Leu Gln Ala
 225 230 235 240
 ACT GAC TCC CAG GGC AAC ACA GTC CTG CAT GCC CTA GTG ATG ATC TCG
 1128
 45 Thr Asp Ser Gln Gly Asn Thr Val Leu His Ala Leu Val Met Ile Ser
 245 250 255
 GAC AAC TCA GCT GAG AAC ATT GCA CTG GTG ACC AGC ATG TAT GAT GGG
 1176
 Asp Asn Ser Ala Glu Asn Ile Ala Leu Val Thr Ser Met Tyr Asp Gly
 260 265 270
 50 CTC CTC CAA GCT GGG GCC CGC CTC TGC CCT ACC GTG CAG CTT GAG GAC
 1224
 55

Leu Leu Gln Ala Gly Ala Arg Leu Cys Pro Thr Val Gln Leu Glu Asp
 275 280 285
 5 ATC CGC AAC CTG CAG GAT CTC ACG CCT CTG AAG CTG GCC GCC AAG GAG
 1272
 Ile Arg Asn Leu Gln Asp Leu Thr Pro Leu Lys Leu Ala Ala Lys Glu
 290 295 300
 10 GGC AAG ATC GAG ATT TTC AGG CAC ATC CTG CAG CGG GAG TTT TCA GGA
 1320
 Gly Lys Ile Glu Ile Phe Arg His Ile Leu Gln Arg Glu Phe Ser Gly
 305 310 315 320
 15 CTG AGC CAC CTT TCC CGA AAG TTC ACC GAG TGG TGC TAT GGG CCT GTC
 1368
 Leu Ser His Leu Ser Arg Lys Phe Thr Glu Trp Cys Tyr Gly Pro Val
 325 330 335
 20 CGG GTG TCG CTG TAT GAC CTG GCT TCT GTG GAC AGC TGT GAG GAG AAC
 1416
 Arg Val Ser Leu Tyr Asp Leu Ala Ser Val Asp Ser Cys Glu Glu Asn
 340 345 350
 25 TCA GTG CTG GAG ATC ATT GCC TTT CAT TGC AAG AGC CCG CAC CGA CAC
 1464
 Ser Val Leu Glu Ile Ile Ala Phe His Cys Lys Ser Pro His Arg His
 355 360 365
 CGA ATG GTC GTT TTG GAG CCC CTG AAC AAA CTG CTG CAG GCG AAA TGG
 1512
 Arg Met Val Val Leu Glu Pro Leu Asn Lys Leu Leu Gln Ala Lys Trp
 370 375 380
 30 GAT CTG CTC ATC CCC AAG TTC TTC TTA AAC TTC CTG TGT AAT CTG ATC
 1560
 Asp Leu Leu Ile Pro Lys Phe Phe Leu Asn Phe Leu Cys Asn Leu Ile
 385 390 395 400
 35 TAC ATG TTC ATC TTC ACC GCT GTT GCC TAC CAT CAG CCT ACC CTG AAG
 1608
 Tyr Met Phe Ile Phe Thr Ala Val Ala Tyr His Gln Pro Thr Leu Lys
 405 410 415
 40 AAG GCC GCC CCT CAC CTG AAA GCG GAG GTT GGA AAC TCC ATG CTG CTG
 1656
 Lys Ala Ala Pro His Leu Lys Ala Glu Val Gly Asn Ser Met Leu Leu
 420 425 430
 45 ACG GGC CAC ATC CTT ATC CTG CTA GGG GGG ATC TAC CTC CTC GTG GGC
 1704
 Thr Gly His Ile Leu Ile Leu Leu Gly Gly Ile Tyr Leu Leu Val Gly
 435 440 445
 50 CAG CTG TGG TAC TTC TGG CGC CGC CAC GTG TTC ATC TGG ATC TCG TTC
 1752
 Gln Leu Trp Tyr Phe Trp Arg Arg His Val Phe Ile Trp Ile Ser Phe
 450 455 460
 ATA GAC AGC TAC TTT GAA ATC CTC TTC CTG TTC CAG GCC CTG CTC ACA
 1800

Ile Asp Ser Tyr Phe Glu Ile Leu Phe Leu Phe Gln Ala Leu Leu Thr
 465 470 475 480
 5 GTG GTG TCC CAG GTG CTG TGT TTC CTG GCC ATC GAG TGG TAC CTG CCC
 1848
 Val Val Ser Gln Val Leu Cys Phe Leu Ala Ile Glu Trp Tyr Leu Pro
 485 490 495
 10 CTG CTT GTG TCT GCG CTG GTG CTG GGC TGG CTG AAC CTG CTT TAC TAT
 1896
 Leu Leu Val Ser Ala Leu Val Leu Gly Trp Leu Asn Leu Leu Tyr Tyr
 500 505 510
 15 ACA CGT GGC TTC CAG CAC ACA GGC ATC TAC AGT GTC ATG ATC CAG AAG
 1944
 Thr Arg Gly Phe Gln His Thr Gly Ile Tyr Ser Val Met Ile Gln Lys
 515 520 525
 20 GTC ATC CTG CCG GAC CTG CTG CGC TTC CTT CTG ATC TAC TTA GTC TTC
 1992
 Val Ile Leu Arg Asp Leu Leu Arg Phe Leu Leu Ile Tyr Leu Val Phe
 530 535 540
 25 CTT TTC GGC TTC GCT GTA GCC CTG GTG AGC CTG AGC CAG GAG GCT TGG
 2040
 Leu Phe Gly Phe Ala Val Ala Leu Val Ser Leu Ser Gln Glu Ala Trp
 545 550 555 560
 CGC CCC GAA GCT CCT ACA GGC CCC AAT GCC ACA GAG TCA GTG CAG CCC
 2088
 Arg Pro Glu Ala Pro Thr Gly Pro Asn Ala Thr Glu Ser Val Gln Pro
 565 570 575
 30 ATG GAG GGA CAG GAG GAC GAG GGC AAC GGG GCC CAG TAC AGG GGT ATC
 2136
 Met Glu Gly Gln Glu Asp Glu Gly Asn Gly Ala Gln Tyr Arg Gly Ile
 580 585 590
 35 CTG GAA GCC TCC TTG GAG CTC TTC AAA TTC ACC ATC GGC ATG GGC GAG
 2184
 Leu Glu Ala Ser Leu Glu Leu Phe Lys Phe Thr Ile Gly Met Gly Glu
 595 600 605
 40 CTG GCC TTC CAG GAG CAG CTG CAC TTC CGC GGC ATG GTG CTG CTG CTG
 2232
 Leu Ala Phe Gln Glu Gln Leu His Phe Arg Gly Met Val Leu Leu Leu
 610 615 620
 45 CTG CTG GCC TAC GTG CTG CTC ACC TAC ATC CTG CTG CTC AAC ATG CTC
 2280
 Leu Leu Ala Tyr Val Leu Leu Thr Tyr Ile Leu Leu Leu Asn Met Leu
 625 630 635 640
 50 ATC GCC CTC ATG AGC GAG ACC GTC AAC AGT GTC GCC ACT GAC AGC TGG
 2328
 Ile Ala Leu Met Ser Glu Thr Val Asn Ser Val Ala Thr Asp Ser Trp
 645 650 655
 AGC ATC TGG AAG CTG CAG AAA GCC ATC TCT GTC CTG GAG ATG GAG AAT
 2376

Ser Ile Trp Lys Leu Gln Lys Ala Ile Ser Val Leu Glu Met Glu Asn
 660 665 670
 5 GGC TAT TGG TGG TGC AGG AAG AAG CAG CGG GCA GGT GTG ATG CTG ACC
 2424
 Gly Tyr Trp Trp Cys Arg Lys Lys Gln Arg Ala Gly Val Met Leu Thr
 675 680 685
 10 GTT GGC ACT AAG CCA GAT GGC AGC CCC GAT GAG CGC TGG TGC TTC AGG
 2472
 Val Gly Thr Lys Pro Asp Gly Ser Pro Asp Glu Arg Trp Cys Phe Arg
 690 695 700
 15 GTG GAG GAG GTG AAC TGG GCT TCA TGG GAG CAG ACG CTG CCT ACG CTG
 2520
 Val Glu Glu Val Asn Trp Ala Ser Trp Glu Gln Thr Leu Pro Thr Leu
 705 710 715 720
 20 TGT GAG GAC CCG TCA GGG GCA GGT GTC CCT CGA ACT CTC GAG AAC CCT
 2568
 Cys Glu Asp Pro Ser Gly Ala Gly Val Pro Arg Thr Leu Glu Asn Pro
 725 730 735
 25 GTC CTG GCT TCC CCT CCC AAG GAG GAT GAG GAT GGT GCC TCT GAG GAA
 2616
 Val Leu Ala Ser Pro Pro Lys Glu Asp Glu Asp Gly Ala Ser Glu Glu
 740 745 750
 AAC TAT GTG CCC GTC CAG CTC CTC CAG TCC AAC TGATGGCCCA GATGCAGCAG
 2669
 Asn Tyr Val Pro Val Gln Leu Leu Gln Ser Asn
 755 760
 30 GAGGCCAGAG GACAGAGCAG AGGATCTTTC CAACCACATC TGCTGGCTCT GGGGTCCCAG
 2729
 TGAATTCTGG TGGCAAATAT ATATTTTCAC TAACTAAAAA AAAAAAAAAA AAAA
 2783
 35

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 763 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 40

(ii) MOLECULE TYPE: protein
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 45

Met Thr Ser Pro Ser Ser Ser Pro Val Phe Arg Leu Glu Thr Leu Asp
 1 5 10 15
 Gly Gly Gln Glu Asp Gly Ser Glu Ala Asp Arg Gly Lys Leu Asp Phe
 20 25 30
 50 Gly Ser Gly Leu Pro Pro Met Glu Ser Gln Phe Gln Gly Glu Asp Arg
 35 40 45
 55

Lys Phe Ala Pro Gln Ile Arg Val Asn Leu Asn Tyr Arg Lys Gly Thr
 50 55 60
 5 Gly Ala Ser Gln Pro Asp Pro Asn Arg Phe Asp Arg Asp Arg Leu Phe
 65 70 75 80
 Asn Ala Val Ser Arg Gly Val Pro Glu Asp Leu Ala Gly Leu Pro Glu
 85 90 95
 10 Tyr Leu Ser Lys Thr Ser Lys Tyr Leu Thr Asp Ser Glu Tyr Thr Glu
 100 105 110
 Gly Ser Thr Gly Lys Thr Cys Leu Met Lys Ala Val Leu Asn Leu Lys
 115 120 125
 15 Asp Gly Val Asn Ala Cys Ile Leu Pro Leu Leu Gln Ile Asp Arg Asp
 130 135 140
 Ser Gly Asn Pro Gln Pro Leu Val Asn Ala Gln Cys Thr Asp Asp Tyr
 145 150 155 160
 20 Tyr Arg Gly His Ser Ala Leu His Ile Ala Ile Glu Lys Arg Ser Leu
 165 170 175
 Gln Cys Val Lys Leu Leu Val Glu Asn Gly Ala Asn Val His Ala Arg
 180 185 190
 25 Ala Cys Gly Arg Phe Phe Gln Lys Gly Gln Gly Thr Cys Phe Tyr Phe
 195 200 205
 Gly Glu Leu Pro Leu Ser Leu Ala Ala Cys Thr Lys Gln Trp Asp Val
 210 215 220
 30 Val Ser Tyr Leu Leu Glu Asn Pro His Gln Pro Ala Ser Leu Gln Ala
 225 230 235 240
 Thr Asp Ser Gln Gly Asn Thr Val Leu His Ala Leu Val Met Ile Ser
 245 250 255
 35 Asp Asn Ser Ala Glu Asn Ile Ala Leu Val Thr Ser Met Tyr Asp Gly
 260 265 270
 Leu Leu Gln Ala Gly Ala Arg Leu Cys Pro Thr Val Gln Leu Glu Asp
 275 280 285
 40 Ile Arg Asn Leu Gln Asp Leu Thr Pro Leu Lys Leu Ala Ala Lys Glu
 290 295 300
 Gly Lys Ile Glu Ile Phe Arg His Ile Leu Gln Arg Glu Phe Ser Gly
 305 310 315 320
 45 Leu Ser His Leu Ser Arg Lys Phe Thr Glu Trp Cys Tyr Gly Pro Val
 325 330 335
 Arg Val Ser Leu Tyr Asp Leu Ala Ser Val Asp Ser Cys Glu Glu Asn
 340 345 350
 50 Ser Val Leu Glu Ile Ile Ala Phe His Cys Lys Ser Pro His Arg His
 355 360 365

55

	Arg Met Val Val Leu Glu Pro Leu Asn Lys Leu Leu Gln Ala Lys Trp	370	375	380
5	Asp Leu Leu Ile Pro Lys Phe Phe Leu Asn Phe Leu Cys Asn Leu Ile	385	390	395 400
	Tyr Met Phe Ile Phe Thr Ala Val Ala Tyr His Gln Pro Thr Leu Lys	405	410	415
10	Lys Ala Ala Pro His Leu Lys Ala Glu Val Gly Asn Ser Met Leu Leu	420	425	430
	Thr Gly His Ile Leu Ile Leu Leu Gly Gly Ile Tyr Leu Leu Val Gly	435	440	445
15	Gln Leu Trp Tyr Phe Trp Arg Arg His Val Phe Ile Trp Ile Ser Phe	450	455	460
	Ile Asp Ser Tyr Phe Glu Ile Leu Phe Leu Phe Gln Ala Leu Leu Thr	465	470	475 480
20	Val Val Ser Gln Val Leu Cys Phe Leu Ala Ile Glu Trp Tyr Leu Pro	485	490	495
	Leu Leu Val Ser Ala Leu Val Leu Gly Trp Leu Asn Leu Leu Tyr Tyr	500	505	510
25	Thr Arg Gly Phe Gln His Thr Gly Ile Tyr Ser Val Met Ile Gln Lys	515	520	525
	Val Ile Leu Arg Asp Leu Leu Arg Phe Leu Leu Ile Tyr Leu Val Phe	530	535	540
30	Leu Phe Gly Phe Ala Val Ala Leu Val Ser Leu Ser Gln Glu Ala Trp	545	550	555 560
	Arg Pro Glu Ala Pro Thr Gly Pro Asn Ala Thr Glu Ser Val Gln Pro	565	570	575
35	Met Glu Gly Gln Glu Asp Glu Gly Asn Gly Ala Gln Tyr Arg Gly Ile	580	585	590
	Leu Glu Ala Ser Leu Glu Leu Phe Lys Phe Thr Ile Gly Met Gly Glu	595	600	605
40	Leu Ala Phe Gln Glu Gln Leu His Phe Arg Gly Met Val Leu Leu Leu	610	615	620
	Leu Leu Ala Tyr Val Leu Leu Thr Tyr Ile Leu Leu Leu Asn Met Leu	625	630	635 640
45	Ile Ala Leu Met Ser Glu Thr Val Asn Ser Val Ala Thr Asp Ser Trp	645	650	655
	Ser Ile Trp Lys Leu Gln Lys Ala Ile Ser Val Leu Glu Met Glu Asn	660	665	670
50	Gly Tyr Trp Trp Cys Arg Lys Lys Gln Arg Ala Gly Val Met Leu Thr	675	680	685

55

Val Gly Thr Lys Pro Asp Gly Ser Pro Asp Glu Arg Trp Cys Phe Arg
 690 695 700
 5 Val Glu Glu Val Asn Trp Ala Ser Trp Glu Gln Thr L u Pro Thr Leu
 705 710 715 720
 Cys Glu Asp Pro Ser Gly Ala Gly Val Pro Arg Thr Leu Glu Asn Pro
 10 725 730 735
 Val Leu Ala Ser Pro Pro Lys Glu Asp Glu Asp Gly Ala Ser Glu Glu
 740 745 750
 15 Asn Tyr Val Pro Val Gln Leu Leu Gln Ser Asn
 755 760

20

Claims

1. An isolated polynucleotide comprising a nucleotide sequence that has at least 80% identity to a nucleotide sequence encoding the hVRCC polypeptide of SEQ ID NO:2 over its entire length; or a nucleotide sequence complementary to said nucleotide sequence.
2. The polynucleotide of claim 1 which is DNA or RNA.
3. A polynucleotide according to one of claims 1 and 2 wherein said nucleotide sequence is at least 80% identical to that contained in SEQ ID NO:1.
4. A polynucleotide according to one of claims 1 to 3 wherein said nucleotide sequence comprises the hVRCC polypeptide encoding sequence contained in SEQ ID NO:1.
5. A polynucleotide according to one of claims 1 to 4 which is polynucleotide of SEQ ID NO: 1.
6. An isolated hVRCC polynucleotide comprising a nucleotide sequence selected from the group consisting of :
 - (a) a nucleotide sequence having at least 80% identity to a nucleotide sequence encoding the hVRCC polypeptide expressed by the cDNA insert deposited at the ATCC with Deposit Number 209625 ; and
 - (b) a nucleotide sequence complementary to the nucleotide sequence of (a).
7. A DNA or RNA molecule comprising an expression system, wherein said expression system is capable of producing a hVRCC polypeptide comprising an amino acid sequence, which has at least 80% identity with the polypeptide of SEQ ID NO:2 when said expression system is present in a compatible host cell.
8. A host cell comprising the expression system of claim 7.
9. A process for producing a hVRCC polypeptide comprising culturing a host of claim 8 under conditions sufficient for the production of said polypeptide and recovering the polypeptide from the culture.
10. A process for producing a cell which produces a hVRCC polypeptide thereof comprising transforming or transfecting a host cell with the expression system of claim 7 such that the host cell, under appropriate culture conditions, produces a hVRCC polypeptide.
11. A hVRCC polypeptide comprising an amino acid sequence which is at least 80% identical to the amino acid sequence of SEQ ID NO:2 over its entire length.

12. The polypeptide of claim 11 which comprises the amino acid sequence of SEQ ID NO:2.
13. An antibody immunospecific for the hVRCC polypeptide of claim 11.
- 5 14. Use of (a) a therapeutically effective amount of an agonist of hVRCC polypeptide of claim 11 and/or (b) a polynucleotide according to one of claims 1 to 6 in a form so as to effect production of said hVRCC polypeptide activity *in vivo*, for the manufacture of a medicament for the treatment of a subject in need of enhanced activity or expression of hVRCC polypeptide.
- 10 15. Use of (a) a therapeutically effective amount of an antagonist of hVRCC polypeptide of claim 11 and/or (b) a nucleic acid molecule that inhibits the expression of the nucleotide sequence encoding said hVRCC polypeptide and/or (c) a therapeutically effective amount of a polypeptide that competes with said hVRCC polypeptide, for the manufacture of a medicament for the treatment of a subject having need to inhibit activity or expression of hVRCC polypeptide.
- 15 16. A process for diagnosing a disease or a susceptibility to a disease in a subject related to expression or activity of hVRCC polypeptide of claim 11 in a subject comprising:
 - (a) determining the presence or absence of a mutation in the nucleotide sequence encoding said hVRCC polypeptide in the genome of said subject; and/or
 - 20 (b) analyzing for the presence or amount of the hVRCC polypeptide expression in a sample derived from said subject.
- 25 17. A method for identifying agonists to hVRCC polypeptide of claim 11 comprising:
 - (a) contacting cells produced by claim 10 with a candidate compound; and
 - (b) determining whether the candidate compound effects a signal generated by activation of the hVRCC polypeptide.
- 30 18. An agonist identified by the method of claim 17.
19. The method for identifying antagonists to hVRCC polypeptide of claim 11 comprising:
 - (a) contacting said cells produced by claim 10 expressing the hVRCC polypeptide on their surface with an agonist;
 - 35 (b) contacting said cells with a candidate compound; and
 - (c) determining whether the signal generated by said agonist is diminished in the presence of the candidate compound.
- 40 20. An antagonist identified by the method of claim 19.



European Patent
Office

PARTIAL EUROPEAN SEARCH REPORT

Application Number

which under Rule 45 of the European Patent Convention EP 98 40 0565 shall be considered, for the purposes of subsequent proceedings, as the European search report

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
D,X	CATERINA, MICHAEL J. ET AL: "The capsaicin receptor: a heat-activated ion channel in the pain pathway" NATURE (LONDON), 1997, 389, 816-824, XP002075020	18,20	C12N15/12 C07K14/705 C07K16/28 A61K38/17 C12Q1/68
Y	* page 822 - page 823; figures 2,5 *	1-6,11, 12	G01N33/50 A61K31/70
X	--- BIRO T ET AL: "Recent advances in understanding of vanilloid receptors: a therapeutic target for treatment of pain and inflammation in skin." J INVESTIG DERMATOL SYMP PROC, AUG 1997, 2 (1) P56-60, XP002075021 UNITED STATES	18,20	
Y	* abstract; figure 1; tables I,II * --- -/--	1-6,11, 12	
			TECHNICAL FIELDS SEARCHED (Int.Cl.6)
			C12N C07K A61K C12Q G01N
INCOMPLETE SEARCH			
<p>The Search Division considers that the present application, or one or more of its claims, does/do not comply with the EPC to such an extent that a meaningful search into the state of the art cannot be carried out, or can only be carried out partially, for these claims.</p> <p>Claims searched completely :</p> <p>Claims searched incompletely :</p> <p>Claims not searched :</p> <p>Reason for the limitation of the search:</p> <p>see sheet C</p>			
Place of search		Date of completion of the search	Examiner
THE HAGUE		20 August 1998	Espen, J
CATEGORY OF CITED DOCUMENTS		<p>T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document</p>	
<p>X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document</p>			

EPO FORM 1503 03/82 (P04C07)



European Patent
Office

PARTIAL EUROPEAN SEARCH REPORT

Application Number
EP 98 40 0565

DOCUMENTS CONSIDERED TO BE RELEVANT			CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	
X	<p>SZALLASI A ET AL: "A novel agonist, phorbol 12-phenylacetate 13-acetate 20-homovanillate, abolishes positive cooperativity of binding by the vanilloid receptor."</p> <p>EUR J PHARMACOL, MAR 28 1996, 299 (1-3) P221-8, XP002059195</p> <p>NETHERLANDS</p> <p>* figure 6 *</p> <p>-----</p>	18,20	
			TECHNICAL FIELDS SEARCHED (Int.Cl.6)

EPO FORM 1503 03.92 (P04C10)



European Patent
Office

INCOMPLETE SEARCH
SHEET C

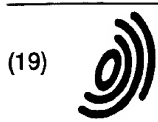
Application Number
EP 98 40 0565

Although claim 16 is directed to a diagnostic method practised on the human/animal body (Article 52(4) EPC), the search has been carried out and based on the alleged effects of the compound/composition.

Claim(s) searched incompletely:
18,20

Reason for the limitation of the search:

Claims 18 and 20 are partially searched, since the subject-matter of said claims is ambiguous and not sufficiently characterized.



Eur päisches Patentamt
Eur pean Patent Office
Office européen des brevets



(11) **EP 0 943 683 A1**

(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:
22.09.1999 Bulletin 1999/38

(21) Application number: **98203620.4**

(22) Date of filing: **27.10.1998**

(51) Int. Cl.⁶: **C12N 15/12, C12N 15/11,
C12N 15/85, C12N 5/10,
C07K 14/705, C12Q 1/68,
G01N 33/50, A61K 31/70**

(84) Designated Contracting States:
**AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE**
Designated Extension States:
AL LT LV MK RO SI

(30) Priority: **10.03.1998 GB 9805137
20.07.1998 GB 9815791
03.09.1998 GB 9819278**

(71) Applicant:
**SMITHKLINE BEECHAM PLC
Brentford, Middlesex TW8 9EP (GB)**

(72) Inventors:
• **Duckworth, David M.**
SmithKline Beecham Pharm.
Harlow, Essex CM19 5AW (GB)

• **Davis, John B.**
SmithKline Beecham Pharm.
Harlow, Essex CM19 5AW (GB)
• **Meadows, Helen J.**
SmithKline Beecham Pharm.
Harlow, Essex CM19 5AW (GB)
• **Hayes, Philip D.**
SmithKline Beecham Pharm.
Harlow, Essex CM19 5AW (GB)

(74) Representative:
Connell, Anthony Christopher et al
SmithKline Beecham plc
Corporate Intellectual Property,
Two New Horizons Court
Brentford, Middlesex TW8 9EP (GB)

(54) **Human vanilloid receptor homologue Vanilrep1**

(57) **VANILREP1 polypeptides and polynucleotides and methods for producing such polypeptides by recombinant techniques are disclosed. Also disclosed are methods for utilizing VANILREP1 polypeptides and polynucleotides in therapy, and diagnostic assays for such.**

EP 0 943 683 A1

Description**Field of the Invention**

5 [0001] This invention relates to newly identified polypeptides and polynucleotides encoding such polypeptides, to their use in therapy and in identifying compounds which may be agonists, antagonists and/or inhibitors which are potentially useful in therapy, and to production of such polypeptides and polynucleotides.

Background of the Invention

10 [0002] The drug discovery process is currently undergoing a fundamental revolution as it embraces 'functional genomics', that is, high throughput genome- or gene-based biology. This approach as a means to identify genes and gene products as therapeutic targets is rapidly superseding earlier approaches based on 'positional cloning'. A phenotype, that is a biological function or genetic disease, would be identified and this would then be tracked back to the responsible gene, based on its genetic map position.

15 [0003] Functional genomics relies heavily on high-throughput DNA sequencing technologies and the various tools of bioinformatics to identify gene sequences of potential interest from the many molecular biology databases now available. There is a continuing need to identify and characterize further genes and their related polypeptides/proteins, as targets for drug discovery.

Summary of the Invention

20 [0004] The present invention relates to VANILREP1, in particular VANILREP1 polypeptides and VANILREP1 polynucleotides, recombinant materials and methods for their production. In another aspect, the invention relates to methods for using such polypeptides and polynucleotides, including the treatment of pain, chronic pain, neuropathic pain, post-operative pain, rheumatoid arthritic pain, neuralgia, neuropathies, algnesia, nerve injury, ischaemia, neurodegeneration, stroke, incontinence and inflammatory disorders hereinafter referred to as "the Diseases", amongst others. In a further aspect, the invention relates to methods for identifying agonists and antagonists/inhibitors using the materials provided by the invention, and treating conditions associated with VANILREP1 imbalance with the identified compounds. In a still further aspect, the invention relates to diagnostic assays for detecting diseases associated with inappropriate VANILREP1 activity or levels.

Description of the Invention

35 [0005] In a first aspect, the present invention relates to VANILREP1 polypeptides. These include the polypeptide of SEQ ID NO:2 and polymorphic variants thereof, for example PVP-1, the polypeptide of SEQ ID NO:8. Such peptides include isolated polypeptides comprising an amino acid sequence which has at least 90% identity, preferably at least 95% identity, more preferably at least 97-99% identity, to that of SEQ ID NO:2 or SEQ ID NO:8 over the entire length of SEQ ID NO:2 or SEQ ID NO:8 respectively. Such polypeptides include those comprising the amino acid of SEQ ID NO:2 or SEQ ID NO:8.

40 [0006] Further peptides of the present invention include isolated polypeptides in which the amino acid sequence has at least 90% identity, preferably at least 95% identity, more preferably at least 97-99% identity, to the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:8, over the entire length of SEQ ID NO:2 or SEQ ID NO:8. Such polypeptides include the polypeptide of SEQ ID NO:2 and SEQ ID NO:8.

45 [0007] Further peptides of the present invention include isolated polypeptides encoded by a polynucleotide comprising the sequence contained in SEQ ID NO:1 or SEQ ID NO:7.

[0008] Polypeptides of the present invention are believed to be members of the ion channel family of polypeptides. They are therefore of interest because they are associated with the mechanism of action of capsaicin (a vanilloid compound), a constituent of chilli peppers. Capsaicin elicits a sensation of burning pain by selectively activating sensory neurons that convey information about noxious stimuli to the central nervous system. The channels are permeable to cations and exhibit a notable preference for divalent cations, particularly calcium ions. The level of calcium ion permeability exceeds that observed for most non-selective cation channels and is similar to values observed for NMDA-type glutamate receptors and $\alpha 7$ nicotinic acetylcholine receptors, both of which are noted for this property. These properties are hereinafter referred to as "VANILREP1 activity" or "VANILREP1 polypeptide activity" or "biological activity of VANILREP1". Also included amongst these activities are antigenic and immunogenic activities of said VANILREP1 polypeptides, in particular the antigenic and immunogenic activities of the polypeptides of SEQ ID NO:2 or SEQ ID NO:8. Preferably, a polypeptide of the present invention exhibits at least one biological activity of VANILREP1.

55 [0009] The polypeptides of the present invention may be in the form of the "mature" protein or may be a part of a larger

protein such as a precursor or a fusion protein. It is often advantageous to include an additional amino acid sequence which contains secretory or leader sequences, prosequences, sequences which aid in purification such as multiple histidine residues, or an additional sequence for stability during recombinant production.

5 [0010] The present invention also includes variants of the aforementioned polypeptides, that is polypeptides that vary from the referents by conservative amino acid substitutions, whereby a residue is substituted by another with like characteristics. Typical such substitutions are among Ala, Val, Leu and Ile; among Ser and Thr; among the acidic residues Asp and Glu; among Asn and Gln; and among the basic residues Lys and Arg; or aromatic residues Phe and Tyr. Particularly preferred are variants in which several, 5-10, 1-5, 1-3, 1-2 or 1 amino acids are substituted, deleted, or added in any combination.

10 [0011] Polypeptides of the present invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

[0012] In a further aspect, the present invention relates to VANILREP1 polynucleotides. Such polynucleotides include 15 isolated polynucleotides comprising a nucleotide sequence encoding a polypeptide which has at least 90% identity, preferably at least 95% identity, to the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:8, over the entire length of SEQ ID NO:2 or SEQ ID NO:8, respectively. In this regard, polypeptides which have at least 97% identity are highly preferred, whilst those with at least 98-99% identity are more highly preferred, and those with at least 99% identity are most highly preferred. Such polynucleotides include a polynucleotide comprising the nucleotide sequence contained in SEQ

20 ID NO:1 or SEQ ID NO:7, encoding the polypeptide of SEQ ID NO:2 or SEQ ID NO:8, respectively.

[0013] Further polynucleotides of the present invention include isolated polynucleotides comprising a nucleotide sequence that has at least 85% identity, preferably at least 90% identity, more preferably at least 95% identity, to a nucleotide sequence encoding a polypeptide of SEQ ID NO:2 or SEQ ID NO:8, respectively, over the entire coding region. In this regard, polynucleotides which have at least 97% identity are highly preferred, whilst those with at least 25 98-99% identity are more highly preferred, and those with at least 99% identity are most highly preferred.

[0014] Further polynucleotides of the present invention include isolated polynucleotides comprising a nucleotide sequence which has at least 85% identity, preferably at least 90% identity, more preferably at least 95% identity, to SEQ ID NO:1 or SEQ ID NO:7, over the entire length of SEQ ID NO:1 or SEQ ID NO:7, respectively. In this regard, polynucleotides which have at least 97% identity are highly preferred, whilst those with at least 98-99% identity are more highly 30 preferred, and those with at least 99% identity are most highly preferred. Such polynucleotides include a polynucleotide comprising the polynucleotide of SEQ ID NO:1 or SEQ ID NO:7, as well as the polynucleotide of SEQ ID NO:1 and SEQ ID NO:7, respectively.

[0015] The invention also provides polynucleotides which are complementary to all the above described polynucleotides.

35 [0016] The nucleotide sequences of SEQ ID NO:1 and SEQ ID NO:7 show homology with rat vanilloid receptor VR1 (M. J. Caterina et al., Nature 389:816-824, 1997). The nucleotide sequence of SEQ ID NO:1 is a cDNA sequence and comprises a polypeptide encoding sequence (nucleotide 864 to 3380) encoding a polypeptide of 839 amino acids, the polypeptide of SEQ ID NO:2. The nucleotide sequence encoding the polypeptide of SEQ ID NO:2 may be identical to the polypeptide encoding sequence contained in SEQ ID NO:1 or it may be a sequence other than the one contained 40 in SEQ ID NO:1, which, as a result of the redundancy (degeneracy) of the genetic code, also encodes the polypeptide of SEQ ID NO:2. The polypeptide of the SEQ ID NO:2 is structurally related to other proteins of the ion channel family, having homology and/or structural similarity with rat vanilloid receptor VR1 (M. J. Caterina et al., Nature 389: 816-824, 1997).

[0017] The nucleotide sequence of SEQ ID NO:7 is a cDNA sequence and comprises a polypeptide encoding 45 sequence (nucleotide 864 to 3380) encoding a polypeptide of 839 amino acids, the polypeptide of SEQ ID NO:8. The nucleotide sequence encoding the polypeptide of SEQ ID NO:8 may be identical to the polypeptide encoding sequence contained in SEQ ID NO:7 or it may be a sequence other than the one contained in SEQ ID NO:7, which, as a result of the redundancy (degeneracy) of the genetic code, also encodes the polypeptide of SEQ ID NO:8. The polypeptide of the SEQ ID NO:8 is structurally related to other proteins of the ion channel family, having homology and/or structural 50 similarity with rat vanilloid receptor VR1 (M. J. Caterina et al., Nature 389: 816-824, 1997).

[0018] Preferred polypeptides and polynucleotides of the present invention are expected to have, *inter alia*, similar biological functions/properties to their homologous polypeptides and polynucleotides. Furthermore, preferred polypeptides and polynucleotides of the present invention have at least one VANILREP1 activity.

55 [0019] The present invention also relates to partial or other polynucleotide and polypeptide sequences which were first identified prior to the determination of the corresponding full length sequences of SEQ ID NO:1, SEQ ID NO:7, SEQ ID NO:2 and SEQ ID NO:8.

[0020] Accordingly, in a further aspect, the present invention provides for an isolated polynucleotide which:

(a) comprises a nucleotide sequence which has at least 85% identity, preferably at least 90% identity, more preferably at least 95% identity, yet more preferably at least 97-99% identity to SEQ ID NO:3 or SEQ ID NO:5, over the entire length of SEQ ID NO:3 or SEQ ID NO:5, respectively;

(b) has a nucleotide sequence which has at least 85% identity, preferably at least 90% identity, more preferably at least 95% identity, yet more preferably at least 97-99% identity, to SEQ ID NO:3 or SEQ ID NO:5, over the entire length of SEQ ID NO:3 or SEQ ID NO:5, respectively;

(c) the polynucleotide of SEQ ID NO:3 or SEQ ID NO:5; or

(d) a nucleotide sequence encoding a polypeptide which has at least 90% identity, preferably at least 95% identity, more preferably at least 97-99% identity, to the amino acid sequence of SEQ ID NO:4 or SEQ ID NO:6, over the entire length of SEQ ID NO:4 or SEQ ID NO:6, respectively;

as well as the polynucleotides of SEQ ID NO:3 and SEQ ID NO:5.

[0021] The present invention further provides for a polypeptide which:

(a) comprises an amino acid sequence which has at least 90% identity, preferably at least 95% identity, more preferably at least 97-99% identity, to that of SEQ ID NO:4 or SEQ ID NO:6, over the entire length of SEQ ID NO:4 or SEQ ID NO:6, respectively;

(b) has an amino acid sequence which has at least 90% identity, preferably at least 95% identity, more preferably at least 97-99% identity, to the amino acid sequence of SEQ ID NO:4 or SEQ ID NO:6, over the entire length of SEQ ID NO:4 or SEQ ID NO:6, respectively;

(c) comprises the amino acid of SEQ ID NO:4 or SEQ ID NO:6; and

(d) is the polypeptide of SEQ ID NO:4 or SEQ ID NO:6;

as well as polypeptides encoded by a polynucleotide comprising the sequence contained in SEQ ID NO:3 or SEQ ID NO:5.

[0022] The nucleotide sequences of SEQ ID NO:3 and SEQ ID NO:5, and the peptide sequences encoded thereby are derived from EST (Expressed Sequence Tag) sequences. It is recognized by those skilled in the art that there will inevitably be some nucleotide sequence reading errors in EST sequences (see Adams, M.D. *et al.*, Nature 377 (supp) 3, 1995). Accordingly, the nucleotide sequences of SEQ ID NO:3 and SEQ ID NO:5, and the peptide sequence encoded therefrom, are therefore subject to the same inherent limitations in sequence accuracy.

[0023] Polynucleotides of the present invention may be obtained, using standard cloning and screening techniques, from a cDNA library derived from mRNA in cells of human brain, cerebellum, dorsal root ganglia, thymus, leukocytes, placenta, foetal liver spleen and ovary, using the expressed sequence tag (EST) analysis (Adams, M.D., *et al.*, Science (1991) 252:1651-1656; Adams, M.D. *et al.*, Nature, (1992) 355:632-634; Adams, M.D., *et al.*, Nature (1995) 377 Supp:3-174). Polynucleotides of the invention can also be obtained from natural sources such as genomic DNA libraries or can be synthesized using well known and commercially available techniques.

[0024] When polynucleotides of the present invention are used for the recombinant production of polypeptides of the present invention, the polynucleotide may include the coding sequence for the mature polypeptide, by itself; or the coding sequence for the mature polypeptide in reading frame with other coding sequences, such as those encoding a leader or secretory sequence, a pre-, or pro- or prepro- protein sequence, or other fusion peptide portions. For example, a marker sequence which facilitates purification of the fused polypeptide can be encoded. In certain preferred embodiments of this aspect of the invention, the marker sequence is a hexa-histidine peptide, as provided in the pQE vector (Qiagen, Inc.) and described in Gentz *et al.*, Proc Natl Acad Sci USA (1989) 86:821-824, or is an HA tag. The polynucleotide may also contain non-coding 5' and 3' sequences, such as transcribed, non-translated sequences, splicing and polyadenylation signals, ribosome binding sites and sequences that stabilize mRNA.

[0025] Further embodiments of the present invention include polynucleotides encoding polypeptide variants which comprise the amino acid sequence of SEQ ID NO:2 or SEQ ID NO:8, respectively and in which several, for instance from 5 to 10, 1 to 5, 1 to 3, 1 to 2 or 1, amino acid residues are substituted, deleted or added, in any combination.

[0026] Polynucleotides which are identical or sufficiently identical to a nucleotide sequence contained in SEQ ID NO:1 or SEQ ID NO:7, respectively, may be used as hybridization probes for cDNA and genomic DNA or as primers for a nucleic acid amplification (PCR) reaction, to isolate full-length cDNAs and genomic clones encoding polypeptides of the present invention and to isolate cDNA and genomic clones of other genes (including genes encoding paralogs from human sources and orthologs and paralogs from species other than human) that have a high sequence similarity to SEQ ID NO:1 or SEQ ID NO:7. Typically these nucleotide sequences are 70% identical, preferably 80% identical, more preferably 90% identical, most preferably 95% identical to that of the referent. The probes or primers will generally comprise at least 15 nucleotides, preferably, at least 30 nucleotides and may have at least 50 nucleotides. Particularly preferred probes will have between 30 and 50 nucleotides. Particularly preferred primers will have between 20 and 25 nucleotides.

[0027] A polynucleotide encoding a polypeptide of the present invention, including homologs from species other than human, may be obtained by a process which comprises the steps of screening an appropriate library under stringent hybridization conditions with a labeled probe having the sequence of SEQ ID NO:1 or SEQ ID NO:7, respectively or a fragment thereof; and isolating full-length cDNA and genomic clones containing said polynucleotide sequence. Such hybridization techniques are well known to the skilled artisan. Preferred stringent hybridization conditions include overnight incubation at 42°C in a solution comprising: 50% formamide, 5xSSC (150mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH7.6), 5x Denhardt's solution, 10 % dextran sulfate, and 20 microgram/ml denatured, sheared salmon sperm DNA; followed by washing the filters in 0.1x SSC at about 65°C. Thus the present invention also includes polynucleotides obtainable by screening an appropriate library under stringent hybridization conditions with a labeled probe having the sequence of SEQ ID NO:1 or SEQ ID NO:7, or a fragment thereof.

[0028] The skilled artisan will appreciate that, in many cases, an isolated cDNA sequence will be incomplete, in that the region coding for the polypeptide is short at the 5' end of the cDNA. This is a consequence of reverse transcriptase, an enzyme with inherently low 'processivity' (a measure of the ability of the enzyme to remain attached to the template during the polymerisation reaction), failing to complete a DNA copy of the mRNA template during 1st strand cDNA synthesis.

[0029] There are several methods available and well known to those skilled in the art to obtain full-length cDNAs, or extend short cDNAs, for example those based on the method of Rapid Amplification of cDNA ends (RACE) (see, for example, Frohman et al., PNAS USA 85, 8998-9002, 1988). Recent modifications of the technique, exemplified by the Marathon™ technology (Clontech Laboratories Inc.) for example, have significantly simplified the search for longer cDNAs. In the Marathon™ technology, cDNAs have been prepared from mRNA extracted from a chosen tissue and an 'adaptor' sequence ligated onto each end. Nucleic acid amplification (PCR) is then carried out to amplify the 'missing' 5' end of the cDNA using a combination of gene specific and adaptor specific oligonucleotide primers. The PCR reaction is then repeated using 'nested' primers, that is, primers designed to anneal within the amplified product (typically an adaptor specific primer that anneals further 3' in the adaptor sequence and a gene specific primer that anneals further 5' in the known gene sequence). The products of this reaction can then be analyzed by DNA sequencing and a full-length cDNA constructed either by joining the product directly to the existing cDNA to give a complete sequence, or carrying out a separate full-length PCR using the new sequence information for the design of the 5' primer.

[0030] Recombinant polypeptides of the present invention may be prepared by processes well known in the art from genetically engineered host cells comprising expression systems. Accordingly, in a further aspect, the present invention relates to expression systems which comprise a polynucleotide or polynucleotides of the present invention, to host cells which are genetically engineered with such expression systems and to the production of polypeptides of the invention by recombinant techniques. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention.

[0031] For recombinant production, host cells can be genetically engineered to incorporate expression systems or portions thereof for polynucleotides of the present invention. Introduction of polynucleotides into host cells can be effected by methods described in many standard laboratory manuals, such as Davis et al, Basic Methods in Molecular Biology (1986) and Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989). Preferred such methods include, for instance, calcium phosphate transfection, DEAE-dextran mediated transfection, transfection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction or infection.

[0032] Representative examples of appropriate hosts include bacterial cells, such as *Streptococci*, *Staphylococci*, *E. coli*, *Streptomyces* and *Bacillus subtilis* cells; fungal cells, such as yeast cells and *Aspergillus* cells; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, HEK 293 and Bowes melanoma cells; and plant cells.

[0033] A great variety of expression systems can be used, for instance, chromosomal, episomal and virus-derived systems, e.g., vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids. The expression systems may contain control regions that regulate as well as engender expression. Generally, any system or vector which is able to maintain, propagate or express a polynucleotide to produce a polypeptide in a host may be used. The appropriate nucleotide sequence may be inserted into an expression system by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook et al., Molecular Cloning, A Laboratory Manual (supra). Appropriate secretion signals may be incorporated into the desired polypeptide to allow secretion of the translated protein into the lumen of the endoplasmic reticulum, the periplasmic space or the extracellular environment. These signals may be endogenous to the polypeptide or they may be heterologous signals.

[0034] If a polypeptide of the present invention is to be expressed for use in screening assays, it is generally preferred that the polypeptide be produced at the surface of the cell. In this event, the cells may be harvested prior to use in the

screening assay. If the polypeptide is secreted into the medium, the medium can be recovered in order to recover and purify the polypeptide. If produced intracellularly, the cells must first be lysed before the polypeptide is recovered.

[0035] Polypeptides of the present invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography is employed for purification. Well known techniques for refolding proteins may be employed to regenerate active conformation when the polypeptide is denatured during intracellular synthesis, isolation and or purification.

[0036] This invention also relates to the use of polynucleotides of the present invention as diagnostic reagents. Detection of a mutated form of the gene characterised by the polynucleotide of SEQ ID NO:1 or SEQ ID NO:7, respectively which is associated with a dysfunction will provide a diagnostic tool that can add to, or define, a diagnosis of a disease, or susceptibility to a disease, which results from under-expression, over-expression or altered spatial or temporal expression of the gene. Individuals carrying mutations in the gene may be detected at the DNA level by a variety of techniques.

[0037] Nucleic acids for diagnosis may be obtained from a subject's cells, such as from blood, urine, saliva, tissue biopsy or autopsy material. The genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR or other amplification techniques prior to analysis. RNA or cDNA may also be used in similar fashion. Deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal genotype. Point mutations can be identified by hybridizing amplified DNA to labeled VANILREP1 nucleotide sequences. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase digestion or by differences in melting temperatures. DNA sequence differences may also be detected by alterations in electrophoretic mobility of DNA fragments in gels, with or without denaturing agents, or by direct DNA sequencing (see, e.g., Myers *et al.*, Science (1985) 230:1242). Sequence changes at specific locations may also be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method (see Cotton *et al.*, Proc Natl Acad Sci USA (1985) 85: 4397-4401). In another embodiment, an array of oligonucleotides probes comprising VANILREP1 nucleotide sequence or fragments thereof can be constructed to conduct efficient screening of e.g., genetic mutations. Array technology methods are well known and have general applicability and can be used to address a variety of questions in molecular genetics including gene expression, genetic linkage, and genetic variability (see for example: M.Chee *et al.*, Science, Vol 274, pp 610-613 (1996)).

[0038] The diagnostic assays offer a process for diagnosing or determining a susceptibility to the Diseases through detection of mutation in the VANILREP1 gene by the methods described. In addition, such diseases may be diagnosed by methods comprising determining from a sample derived from a subject an abnormally decreased or increased level of polypeptide or mRNA. Decreased or increased expression can be measured at the RNA level using any of the methods well known in the art for the quantitation of polynucleotides, such as, for example, nucleic acid amplification, for instance PCR, RT-PCR, RNase protection, Northern blotting and other hybridization methods. Assay techniques that can be used to determine levels of a protein, such as a polypeptide of the present invention, in a sample derived from a host are well-known to those of skill in the art. Such assay methods include radioimmunoassays, competitive-binding assays, Western Blot analysis and ELISA assays.

[0039] Thus in another aspect, the present invention relates to a diagnostic kit which comprises:

- (a) a polynucleotide of the present invention, preferably the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO:7, respectively or a fragment thereof;
- (b) a nucleotide sequence complementary to that of (a);
- (c) a polypeptide of the present invention, preferably the polypeptide of SEQ ID NO:2 or SEQ ID NO:8, respectively or a fragment thereof; or
- (d) an antibody to a polypeptide of the present invention, preferably to the polypeptide of SEQ ID NO:2 or SEQ ID NO:8, respectively.

[0040] It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component. Such a kit will be of use in diagnosing a disease or susceptibility to a disease, particularly pain, chronic pain, neuropathic pain, postoperative pain, rheumatoid arthritic pain, neuralgia, neuropathies, algosia, nerve injury, ischaemia, neurodegeneration, stroke, incontinence and inflammatory disorders, amongst others.

[0041] The nucleotide sequences of the present invention are also valuable for chromosome localisation. The sequence is specifically targeted to, and can hybridize with, a particular location on an individual human chromosome. The mapping of relevant sequences to chromosomes according to the present invention is an important first step in correlating those sequences with gene associated disease. Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found in, for example, V. McKusick, Mendelian Inheritance in Man (available on-line through Johns Hopkins Univer-

sity Welch Medical Library). The relationship between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (coinheritance of physically adjacent genes).

[0042] The differences in the cDNA or genomic sequence between affected and unaffected individuals can also be determined. If a mutation is observed in some or all of the affected individuals but not in any normal individuals, then

the mutation is likely to be the causative agent of the disease.

[0043] The gene of the present invention maps to human chromosome 17p13. The nucleotide sequences of the present invention are also valuable for tissue localisation. Such techniques allow the determination of expression patterns of the human VANILREP1 polypeptides in tissues by detection of the mRNAs that encode them. These techniques include in situ hybridization techniques and nucleotide amplification techniques, for example PCR. Such techniques are well known in the art. Results from these studies provide an indication of the normal functions of the polypeptides in the organism. In addition, comparative studies of the normal expression pattern of human VANILREP1 mRNAs with that of mRNAs encoded by a human VANILREP1 gene provide valuable insights into the role of mutant human VANILREP1 polypeptides, or that of inappropriate expression of normal human VANILREP1 polypeptides, in disease. Such inappropriate expression may be of a temporal, spatial or simply quantitative nature.

[0044] The polypeptides of the invention or their fragments or analogs thereof, or cells expressing them, can also be used as immunogens to produce antibodies immunospecific for polypeptides of the present invention. The term "immunospecific" means that the antibodies have substantially greater affinity for the polypeptides of the invention than their affinity for other related polypeptides in the prior art.

[0045] Antibodies generated against polypeptides of the present invention may be obtained by administering the polypeptides or epitope-bearing fragments, analogs or cells to an animal, preferably a non-human animal, using routine protocols. For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler, G. and Milstein, C., *Nature* (1975) 256:495-497), the trioma technique, the human B-cell hybridoma technique (Kozbor *et al.*, *Immunology Today* (1983) 4:72) and the EBV-hybridoma technique (Cole *et al.*, *Monoclonal Antibodies and Cancer Therapy*, 77-96, Alan R. Liss, Inc., 1985).

[0046] Techniques for the production of single chain antibodies, such as those described in U.S. Patent No. 4,946,778, can also be adapted to produce single chain antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms, including other mammals, may be used to express humanized antibodies.

[0047] The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptide or to purify the polypeptides by affinity chromatography.

[0048] Antibodies against polypeptides of the present invention may also be employed to treat the Diseases, amongst others.

[0049] In a further aspect, the present invention relates to genetically engineered soluble fusion proteins comprising a polypeptide of the present invention, or a fragment thereof and various portions of the constant regions of heavy or light chains of immunoglobulins of various subclasses (IgG, IgM, IgA, IgE). Preferred as an immunoglobulin is the constant part of the heavy chain of human IgG, particularly IgG1, where fusion takes place at the hinge region. In a particular embodiment, the Fc part can be removed simply by incorporation of a cleavage sequence which can be cleaved with blood clotting factor Xa. Furthermore, this invention relates to processes for the preparation of these fusion proteins by genetic engineering, and to the use thereof for drug screening, diagnosis and therapy. A further aspect of the invention also relates to polynucleotides encoding such fusion proteins. Examples of fusion protein technology can be found in International Patent Application Nos. WO94/29458 and WO94/22914.

[0050] Another aspect of the invention relates to a method for inducing an immunological response in a mammal which comprises inoculating the mammal with a polypeptide of the present invention, adequate to produce antibody and/or T cell immune response to protect said animal from the Diseases hereinbefore mentioned amongst others. Yet another aspect of the invention relates to a method of inducing immunological response in a mammal which comprises, delivering a polypeptide of the present invention via a vector directing expression of the polynucleotide and coding for the polypeptide *in vivo* in order to induce such an immunological response to produce antibody to protect said animal from diseases.

[0051] A further aspect of the invention relates to an immunological/vaccine formulation (composition) which, when introduced into a mammalian host, induces an immunological response in that mammal to a polypeptide of the present invention wherein the composition comprises a polypeptide or polynucleotide of the present invention. The vaccine formulation may further comprise a suitable carrier. Since a polypeptide may be broken down in the stomach, it is preferably administered parenterally (for instance, subcutaneous intramuscular, intravenous, or intradermal injection). Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation isotonic with the blood of the recipient; and aqueous and non-aqueous sterile suspensions which may include suspending agents or thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example, sealed ampoules and vials and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately

prior to use. The vaccine formulation may also include adjuvant systems for enhancing the immunogenicity of the formulation, such as oil-in water systems and other systems known in the art. The dosage will depend on the specific activity of the vaccine and can be readily determined by routine experimentation.

[0052] Polypeptides of the present invention are responsible for one or more biological functions, including one or more disease states, in particular the Diseases hereinbefore mentioned. It is therefore desirable to devise screening methods to identify compounds which stimulate or which inhibit the function of the polypeptide. Accordingly, in a further aspect, the present invention provides for a method of screening compounds to identify those which stimulate or which inhibit the function of the polypeptide. In general, agonists or antagonists may be employed for therapeutic and prophylactic purposes for such Diseases as hereinbefore mentioned. Compounds may be identified from a variety of sources, for example, cells, cell-free preparations, chemical libraries, and natural product mixtures. Such agonists, antagonists or inhibitors so-identified may be natural or modified substrates, ligands, receptors, enzymes, etc., as the case may be, of the polypeptide; or may be structural or functional mimetics thereof (see Coligan *et al.*, Current Protocols in Immunology 1(2):Chapter 5 (1991)).

[0053] The screening method may simply measure the binding of a candidate compound to the polypeptide, or to cells or membranes bearing the polypeptide, or a fusion protein thereof by means of a label directly or indirectly associated with the candidate compound. Alternatively, the screening method may involve competition with a labeled competitor. Further, these screening methods may test whether the candidate compound results in a signal generated by activation or inhibition of the polypeptide, using detection systems appropriate to the cells bearing the polypeptide. Inhibitors of activation are generally assayed in the presence of a known agonist and the effect on activation by the agonist by the presence of the candidate compound is observed. Constitutively active polypeptides may be employed in screening methods for inverse agonists or inhibitors, in the absence of an agonist or inhibitor, by testing whether the candidate compound results in inhibition of activation of the polypeptide. Further, the screening methods may simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide of the present invention, to form a mixture, measuring VANILREP1 activity in the mixture, and comparing the VANILREP1 activity of the mixture to a standard.

Fusion proteins, such as those made from Fc portion and VANILREP1 polypeptide, as hereinbefore described, can also be used for high-throughput screening assays to identify antagonists for the polypeptide of the present invention (see D. Bennett *et al.*, J Mol Recognition, 8:52-58 (1995); and K. Johanson *et al.*, J Biol Chem, 270(16):9459-9471 (1995)).

[0054] The polynucleotides, polypeptides and antibodies to the polypeptide of the present invention may also be used to configure screening methods for detecting the effect of added compounds on the production of mRNA and polypeptide in cells. For example, an ELISA assay may be constructed for measuring secreted or cell associated levels of polypeptide using monoclonal and polyclonal antibodies by standard methods known in the art. This can be used to discover agents which may inhibit or enhance the production of polypeptide (also called antagonist or agonist, respectively) from suitably manipulated cells or tissues.

[0055] The polypeptide may be used to identify membrane bound or soluble receptors, if any, through standard receptor binding techniques known in the art. These include, but are not limited to, ligand binding and crosslinking assays in which the polypeptide is labeled with a radioactive isotope (for instance, ^{125}I), chemically modified (for instance, biotinylated), or fused to a peptide sequence suitable for detection or purification, and incubated with a source of the putative receptor (cells, cell membranes, cell supernatants, tissue extracts, bodily fluids). Other methods include biophysical techniques such as surface plasmon resonance and spectroscopy. These screening methods may also be used to identify agonists and antagonists of the polypeptide which compete with the binding of the polypeptide to its receptors, if any. Standard methods for conducting such assays are well understood in the art.

[0056] Examples of potential polypeptide antagonists include antibodies or, in some cases, oligonucleotides or proteins which are closely related to the ligands, substrates, receptors, enzymes, etc., as the case may be, of the polypeptide, e.g., a fragment of the ligands, substrates, receptors, enzymes, etc.; or small molecules which bind to the polypeptide of the present invention but do not elicit a response, so that the activity of the polypeptide is prevented.

[0057] Thus, in another aspect, the present invention relates to a screening kit for identifying agonists, antagonists, ligands, receptors, substrates, enzymes, etc. for polypeptides of the present invention; or compounds which decrease or enhance the production of such polypeptides, which comprises:

- (a) a polypeptide of the present invention;
 - (b) a recombinant cell expressing a polypeptide of the present invention;
 - (c) a cell membrane expressing a polypeptide of the present invention; or
 - (d) antibody to a polypeptide of the present invention;
- which polypeptide is preferably that of SEQ ID NO:2 or SEQ ID NO:8.

[0058] It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component.

[0059] It will be readily appreciated by the skilled artisan that a polypeptide of the present invention may also be used in a method for the structure-based design of an agonist, antagonist or inhibitor of the polypeptide, by:

- (a) determining in the first instance the three-dimensional structure of the polypeptide;
- (b) deducing the three-dimensional structure for the likely reactive or binding site(s) of an agonist, antagonist or inhibitor;
- (c) synthesising candidate compounds that are predicted to bind to or react with the deduced binding or reactive site; and
- (d) testing whether the candidate compounds are indeed agonists, antagonists or inhibitors.

It will be further appreciated that this will normally be an iterative process.

[0060] In a further aspect, the present invention provides methods of treating abnormal conditions such as, for instance, pain, chronic pain, neuropathic pain, postoperative pain, rheumatoid arthritic pain, neuralgia, neuropathies, algesia, nerve injury, ischaemia, neurodegeneration, stroke, incontinence and inflammatory disorders, related to either an excess of, or an under-expression of, VANILREP1 polypeptide activity.

[0061] If the activity of the polypeptide is in excess, several approaches are available. One approach comprises administering to a subject in need thereof an inhibitor compound (antagonist) as hereinabove described, optionally in combination with a pharmaceutically acceptable carrier, in an amount effective to inhibit the function of the polypeptide, such as, for example, by blocking the binding of ligands, substrates, receptors, enzymes, etc., or by inhibiting a second signal, and thereby alleviating the abnormal condition. In another approach, soluble forms of the polypeptides still capable of binding the ligand, substrate, enzymes, receptors, etc. in competition with endogenous polypeptide may be administered. Typical examples of such competitors include fragments of the VANILREP1 polypeptide.

[0062] In still another approach, expression of the gene encoding endogenous VANILREP1 polypeptide can be inhibited using expression blocking techniques. Known such techniques involve the use of antisense sequences, either internally generated or externally administered (see, for example, O'Connor, J Neurochem (1991) 56:560 in Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Alternatively, oligonucleotides which form triple helices ("triplexes") with the gene can be supplied (see, for example, Lee *et al.*, Nucleic Acids Res (1979) 6:3073; Cooney *et al.*, Science (1988) 241:456; Dervan *et al.*, Science (1991) 251:1360). These oligomers can be administered *per se* or the relevant oligomers can be expressed *in vivo*. Synthetic antisense or triplex oligonucleotides may comprise modified bases or modified backbones. Examples of the latter include methylphosphonate, phosphorothioate or peptide nucleic acid backbones. Such backbones are incorporated in the antisense or triplex oligonucleotide in order to provide protection from degradation by nucleases and are well known in the art. Antisense and triplex molecules synthesised with these or other modified backbones also form part of the present invention.

[0063] In addition, expression of the human VANILREP1 polypeptide may be prevented by using ribozymes specific to the human VANILREP1 mRNA sequence. Ribozymes are catalytically active RNAs that can be natural or synthetic (see for example Usman, N, *et al.*, Curr. Opin. Struct. Biol (1996) 6(4), 527-33.) Synthetic ribozymes can be designed to specifically cleave human VANILREP1 mRNAs at selected positions thereby preventing translation of the human VANILREP1 mRNAs into functional polypeptide. Ribozymes may be synthesised with a natural ribose phosphate backbone and natural bases, as normally found in RNA molecules. Alternatively the ribozymes may be synthesised with non-natural backbones to provide protection from ribonuclease degradation, for example, 2'-O-methyl RNA, and may contain modified bases.

[0064] For treating abnormal conditions related to an under-expression of VANILREP1 and its activity, several approaches are also available. One approach comprises administering to a subject a therapeutically effective amount of a compound which activates a polypeptide of the present invention, i.e., an agonist as described above, in combination with a pharmaceutically acceptable carrier, to thereby alleviate the abnormal condition. Alternatively, gene therapy may be employed to effect the endogenous production of VANILREP1 by the relevant cells in the subject. For example, a polynucleotide of the invention may be engineered for expression in a replication defective retroviral vector, as discussed above. The retroviral expression construct may then be isolated and introduced into a packaging cell transduced with a retroviral plasmid vector containing RNA encoding a polypeptide of the present invention such that the packaging cell now produces infectious viral particles containing the gene of interest. These producer cells may be administered to a subject for engineering cells *in vivo* and expression of the polypeptide *in vivo*. For an overview of gene therapy, see Chapter 20, Gene Therapy and other Molecular Genetic-based Therapeutic Approaches, (and references cited therein) in Human Molecular Genetics, T Strachan and A P Read, BIOS Scientific Publishers Ltd (1996). Another approach is to administer a therapeutic amount of a polypeptide of the present invention in combination with a suitable pharmaceutical carrier.

[0065] In a further aspect, the present invention provides for pharmaceutical compositions comprising a therapeutically effective amount of a polypeptide, such as the soluble form of a polypeptide of the present invention, agonist/antagonist peptide or small molecule compound, in combination with a pharmaceutically acceptable carrier or excipient. Such carriers include, but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol, and combinations thereof. The invention further relates to pharmaceutical packs and kits comprising one or more containers

filled with one or more of the ingredients of the aforementioned compositions of the invention. Polypeptides and other compounds of the present invention may be employed alone or in conjunction with other compounds, such as therapeutic compounds.

[0066] The composition will be adapted to the route of administration, for instance by a systemic or an oral route. Preferred forms of systemic administration include injection, typically by intravenous injection. Other injection routes, such as subcutaneous, intramuscular, or intraperitoneal, can be used. Alternative means for systemic administration include transmucosal and transdermal administration using penetrants such as bile salts or fusidic acids or other detergents. In addition, if a polypeptide or other compounds of the present invention can be formulated in an enteric or an encapsulated formulation, oral administration may also be possible. Administration of these compounds may also be topical and/or localized, in the form of salves, pastes, gels, and the like.

[0067] The dosage range required depends on the choice of peptide or other compounds of the present invention, the route of administration, the nature of the formulation, the nature of the subject's condition, and the judgment of the attending practitioner. Suitable dosages, however, are in the range of 0.1-100 µg/kg of subject. Wide variations in the needed dosage, however, are to be expected in view of the variety of compounds available and the differing efficiencies of various routes of administration. For example, oral administration would be expected to require higher dosages than administration by intravenous injection. Variations in these dosage levels can be adjusted using standard empirical routines for optimization, as is well understood in the art.

[0068] Polypeptides used in treatment can also be generated endogenously in the subject, in treatment modalities often referred to as "gene therapy" as described above. Thus, for example, cells from a subject may be engineered with a polynucleotide, such as a DNA or RNA, to encode a polypeptide *ex vivo*, and for example, by the use of a retroviral plasmid vector. The cells are then introduced into the subject.

[0069] Polynucleotide and polypeptide sequences form a valuable information resource with which to identify further sequences of similar homology. This is most easily facilitated by storing the sequence in a computer readable medium and then using the stored data to search a sequence database using well known searching tools, such as those in the GCG and Lasergene software packages. Accordingly, in a further aspect, the present invention provides for a computer readable medium having stored thereon a polynucleotide comprising the sequence of SEQ ID NO:1 or SEQ ID NO:7 and/or a polypeptide sequence encoded thereby.

[0070] The following definitions are provided to facilitate understanding of certain terms used frequently hereinbefore.

[0071] "Antibodies" as used herein includes polyclonal and monoclonal antibodies, chimeric, single chain, and humanized antibodies, as well as Fab fragments, including the products of an Fab or other immunoglobulin expression library.

[0072] "Isolated" means altered "by the hand of man" from the natural state. If an "isolated" composition or substance occurs in nature, it has been changed or removed from its original environment, or both. For example, a polynucleotide or a polypeptide naturally present in a living animal is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein.

[0073] "Polynucleotide" generally refers to any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. "Polynucleotides" include, without limitation, single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, "polynucleotide" refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The term "polynucleotide" also includes DNAs or RNAs containing one or more modified bases and DNAs or RNAs with backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications may be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically or metabolically modified forms of polynucleotides as typically found in nature, as well as the chemical forms of DNA and RNA characteristic of viruses and cells. "Polynucleotide" also embraces relatively short polynucleotides, often referred to as oligonucleotides.

[0074] "Polypeptide" refers to any peptide or protein comprising two or more amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres. "Polypeptide" refers to both short chains, commonly referred to as peptides, oligopeptides or oligomers, and to longer chains, generally referred to as proteins. Polypeptides may contain amino acids other than the 20 gene-encoded amino acids. "Polypeptides" include amino acid sequences modified either by natural processes, such as post-translational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications may occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present to the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched and branched cyclic polypeptides may result from post-

translation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, biotinylation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cystine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination (see, for instance, *Proteins - Structure and Molecular Properties*, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York, 1993; Wold, F., *Post-translational Protein Modifications: Perspectives and Prospects*, pgs. 1-12 in *Post-translational Covalent Modification of Proteins*, B. C. Johnson, Ed., Academic Press, New York, 1983; Seifter *et al.*, "Analysis for protein modifications and nonprotein cofactors", *Meth Enzymol* (1990) 182:626-646 and Rattan *et al.*, "Protein Synthesis: Post-translational Modifications and Aging", *Ann NY Acad Sci* (1992) 663:48-62).

[0075] "Variant" refers to a polynucleotide or polypeptide that differs from a reference polynucleotide or polypeptide, but retains essential properties. A typical variant of a polynucleotide differs in nucleotide sequence from another, reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. A typical variant of a polypeptide differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. A variant of a polynucleotide or polypeptide may be a naturally occurring such as an allelic variant, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of polynucleotides and polypeptides may be made by mutagenesis techniques or by direct synthesis.

[0076] "Identity," as known in the art, is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the match between strings of such sequences. "Identity" and "similarity" can be readily calculated by known methods, including but not limited to those described in (*Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data*, Part I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987; and *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; and Carillo, H., and Lipman, D., *SIAM J. Applied Math.*, 48: 1073 (1988). Preferred methods to determine identity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in publicly available computer programs. Preferred computer program methods to determine identity and similarity between two sequences include, but are not limited to, the GCG program package (Devereux, J., et al., *Nucleic Acids Research* 12(1): 387 (1984)), BLASTP, BLASTN, and FASTA (Altschul, S.F. et al., *J. Molec. Biol.* 215: 403-410 (1990)). The BLAST X program is publicly available from NCBI and other sources (*BLAST Manual*, Altschul, S., et al., NCBI NLM NIH Bethesda, MD 20894; Altschul, S., et al., *J. Mol. Biol.* 215: 403-410 (1990)). The well known Smith Waterman algorithm may also be used to determine identity.

[0077] Preferred parameters for polypeptide sequence comparison include the following:

- 1) Algorithm: Needleman and Wunsch, *J. Mol Biol.* 48: 443-453 (1970)
Comparison matrix: BLOSSUM62 from Hentikoff and Hentikoff, *Proc. Nat. Acad. Sci. USA.* 89:10915-10919 (1992)
Gap Penalty: 12
Gap Length Penalty: 4

[0078] A program useful with these parameters is publicly available as the "gap" program from Genetics Computer Group, Madison WI. The aforementioned parameters are the default parameters for peptide comparisons (along with no penalty for end gaps).

[0079] Preferred parameters for polynucleotide comparison include the following:

- 1) Algorithm: Needleman and Wunsch, *J. Mol Biol.* 48: 443-453 (1970)
Comparison matrix: matches = +10, mismatch = 0
Gap Penalty: 50
Gap Length Penalty: 3

Available as: The "gap" program from Genetics Computer Group, Madison WI. These are the default parameters for nucleic acid comparisons.

[0080] By way of example, a polynucleotide sequence of the present invention may be identical to the reference sequence of SEQ ID NO:1, that is be 100% identical, or it may include up to a certain integer number of nucleotide alterations as compared to the reference sequence. Such alterations are selected from the group consisting of at least one nucleotide deletion, substitution, including transition and transversion, or insertion, and wherein said alterations may occur at the 5' or 3' terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among the nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence. The number of nucleotide alterations is determined by multiplying the total number of nucleotides in SEQ ID NO:1 by the numerical percent of the respective percent identity (divided by 100) and subtracting that product from said total number of nucleotides in SEQ ID NO:1, or:

$$n_n \leq x_n - (x_n \cdot y),$$

wherein n_n is the number of nucleotide alterations, x_n is the total number of nucleotides in SEQ ID NO:1, and y is, for instance, 0.70 for 70%, 0.80 for 80%, 0.85 for 85%, 0.90 for 90%, 0.95 for 95%, etc., and wherein any non-integer product of x_n and y is rounded down to the nearest integer prior to subtracting it from x_n . Alterations of a polynucleotide sequence encoding the polypeptide of SEQ ID NO:2 may create nonsense, missense or frameshift mutations in this coding sequence and thereby alter the polypeptide encoded by the polynucleotide following such alterations.

[0081] Similarly, a polypeptide sequence of the present invention may be identical to the reference sequence of SEQ ID NO:2, that is be 100% identical, or it may include up to a certain integer number of amino acid alterations as compared to the reference sequence such that the % identity is less than 100%. Such alterations are selected from the group consisting of at least one amino acid deletion, substitution, including conservative and non-conservative substitution, or insertion, and wherein said alterations may occur at the amino- or carboxy-terminal positions of the reference polypeptide sequence or anywhere between those terminal positions, interspersed either individually among the amino acids in the reference sequence or in one or more contiguous groups within the reference sequence. The number of amino acid alterations for a given % identity is determined by multiplying the total number of amino acids in SEQ ID NO:2 by the numerical percent of the respective percent identity (divided by 100) and then subtracting that product from said total number of amino acids in SEQ ID NO:2, or:

$$n_a \leq x_a - (x_a \cdot y),$$

wherein n_a is the number of amino acid alterations, x_a is the total number of amino acids in SEQ ID NO:2, and y is, for instance 0.70 for 70%, 0.80 for 80%, 0.85 for 85% etc., and wherein any non-integer product of x_a and y is rounded down to the nearest integer prior to subtracting it from x_a .

[0082] "Homolog" is a generic term used in the art to indicate a polynucleotide or polypeptide sequence possessing a high degree of sequence relatedness to a subject sequence. Such relatedness may be quantified by determining the degree of identity and/or similarity between the sequences being compared as hereinbefore described. Falling within this generic term are the terms "ortholog", meaning a polynucleotide or polypeptide that is the functional equivalent of a polynucleotide or polypeptide in another species, and "paralog" meaning a functionally similar sequence when considered within the same species.

[0083] "Fusion protein" refers to a protein encoded by two, often unrelated, fused genes or fragments thereof. In one example, EP-A-0 464 discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, employing an immunoglobulin Fc region as a part of a fusion protein is advantageous for use in therapy and diagnosis resulting in, for example, improved pharmacokinetic properties [see, e.g., EP-A 0232 262]. On the other hand, for some uses it would be desirable to be able to delete the Fc part after the fusion protein has been expressed, detected and purified.

[0084] All publications, including but not limited to patents and patent applications, cited in this specification are herein incorporated by reference as if each individual publication were specifically and individually indicated to be incorporated by reference herein as though fully set forth.

Example**Example 1 - Electrophysiological Studies**

5 [0085] *Xenopus laevis* oocyte removal and dissociation were performed. Injections of cDNA for VANILREP1 were made into the nuclei of defolliculated oocytes (1.5ng cDNA/oocyte). After injection the oocytes were incubated between 19-22°C in modified Barth's solution (MBS) plus gentamycin (0.1mg/ml, pH 7.4) and used for electrophysiological recordings within 2-4 days.

10 [0086] For electrophysiological recordings oocytes were placed in a recording chamber and continuously perfused with a solution containing in mM: NaCl 88, KCl 1, NaHCO₃ 2.4, HEPES 15, MgCl₂ 1, BaCl₂ 0.1. (14 ml min⁻¹). Solution was applied using large bore tubing (internal diameter 1.5mm) which facilitated rapid solution exchange (half-time 350 - 1000 ms). Oocytes were held under voltage clamp at -60m V using the two-electrode voltage-clamp technique. Electrodes were low resistance (0.5-3 MΩ) and were filled with 3 M KCl.

15 [0087] Currents were evoked in response to application of capsaicin applied through the perfusion system until the maximum current amplitude was reached. Figure 1 shows currents evoked from a cell in response to capsaicin over a range of concentrations. Figure 2 shows inhibition of the capsaicin evoked response (1 μM) by capsazepine (10μM). The inhibition is reversible with extended washing.

20

25

30

35

40

45

50

55

SEQUENCE INFORMATION

SEQ ID NO:1

5 CAGCGTCGGGTGCAGTTTGGCCGGAGGTTGCAGTGAGCAGAGATTGCCCCATTGCACTCT
 AGTCTGGGCGACAGGGTGAGACACACACACACACACACACACACACACACACAC
 ACACAAGCCTAAACATTCRAGGCCAGGATGCTTGACAGATGTTGATTCATAAAATGACA
 AAAAGCACAAAATCCAAAATCTCGTATAAGCTCAGTGGCTGTGGCAGCGAGGTTGAAGAG
 10 CAAAGGCAGGCCGGGCACCTGGCTGATGATGTGTGGACCCGTTGCACAGCAGGGCCCCGC
 AGTGGGTGTGGGTGTGGGTGTGGGTGGGCCAGTYTCTGCCGCTCACCTATTCCAGGGA
 CACAGTCTGCTTGGCTCTTCTGGACTGAGCCATCCTCATCACCGAGATCCTCCCTGAATT
 CAGCCCCACGACAGCCACCCCGGCCGTTTTCTTGTCTGTGTGGGGAGGGAGGCAGCGCG
 15 GTGGTTATCAACCTCACCTGCAGAGGAGGCACCTGAGGCCAGAGACGAGGAGGATGG
 GTCTAACCAGAACACAGATGGCTCTGAGCCGGGGCCTGTCCACCTCCAGGCCGAC
 GTCAGTGGCCGACGAGCTGCTGGGCCCTGCTAGGCCCTGCTCACCTCTGAGGCCTCTGGG
 GTGAGAGGTTCACTCTTGAAACACTTCAGTTCTAGGGGCTGGGGCAGCAGCAAGTTG
 GAGTTTTGGGGTACCCTGCTTCACAGGGCCCTTGGCAAGGAGGGCAGGTGGGGTCTAAGG
 20 ACAAGCAGTCTTACTTTGGGAGTCAACCCGGCGTGGTGGCTGCTGCAGGTGGCACACT
 GGGCCACAGAGGATCCAGCAAGGATGAAGAAATGGAGCAGCACAGACTTGGGGGCAGCTG
 CGGACCCACTCCAAAGGACACCTGCCCGAGACCCCTGGATGGAGACCTAACTCCAGGC
 CACCTCCAGCCAAGCCCCAGCTCTCCACGGCCAAGAGCCGACCCGGCTCTTTGGGAAGG
 25 GTGACTCGGAGGAGGCTTTCCCGGTGGATTGCCCTCACGAGGAAGGTGAGCTGGACTCCT
 GCCCCACCATCACAGTCAGCCCTGTTATCACCATCCAGAGGCCAGGAGACGGCCCCACCG
 GTGCCAGGCTGCTGTCCAGGACTCTGTGCGCCGACGACCGAGAAGACCCTCAGGCTCT
 ATGATCGCAGGAGTATCTTTGAAGCCGTTGCTCAGAATAACTGCCAGGATCTGGAGAGCC
 30 TGCTGCTCTTCTGCAGAAGAGCAAGAAGCACCTCACAGACAACGAGTTCAAAGACCCTG
 AGACAGGGAAGACCTGTCTGCTGAAAGCCATGCTCAACCTGCACGACGGACAGAACACCA
 CCATCCCCCTGCTCTGGAGATCGCGCGCAACGGACAGCCTGAAGGAGCTTGTCAACG
 CCAGCTACACGGACAGCTACTACAAGGGCCAGACAGCACTGCACATCGCCATCGAGAGAC
 35 GCAACATGGCCCTGGTGACCCTCTGTGTGGAGAACGGAGCAGACGTCAGGCTGCGGCC
 ATGGGGACTTCTTTAAGAAAACCAAAGGGCGGCTGGATTCTACTTCGGTGAACTGCCCC
 TGTCCCTGGCCGCGTGACCAACCAGCTGGGCATCGTGAAGTTCTGTCTGCAGAACTCCT
 GGCAGACGGCCGACATCAGCGCCAGGGACTCGGTGGGCAACACGGTGCTGCACGCCCTGG
 40 TGGAGGTGGCCGACAACACGGCCGACAACACGAAGTTTGTGACGAGCATGTACAATGAGA
 TTCTGATCCTGGGGCCAACTGCACCCGACGCTGAAGCTGGAGGAGCTCACCACAAGA
 AGGGAATGACCCGCTGGCTCTGGCAGCTGGGACCGGAAGATCGGGTCTTGGCCTATA
 TTCTCCAGCGGGAGATCCAGGAGCCGAGTGACGGCACCTGTCCAGGAAGTTCACCGAGT
 45 GGGCCTACGGGCCCGTGCACTCTCGCTGTACGACCTGTCTGATCGACACCTGCGAGA
 AGAACTCGGTGCTGGAGGTGATCGCTACAGCAGCAGCGAGACCCCTAATCGCCACGACA
 TGCTCTTGGTGGAGCCGCTGAACCGACTCCTGCAGACAAGTGGGACAGATTCTGTCAAGC
 GCATCTTCTACTTCAACTTCTGTGCTACTGCTGTACATGATCATCTTACCATGGCTG
 50 CCTACTACAGGCCCGTGGATGGCTTGCTCCCTTTAAGATGGAAAAAAGTGGAGACTATT
 TCCGAGTTACTGGAGAGATCCTGTCTGTGTTAGGAGGAGTCTACTTCTTTTCCGAGGGA
 TTCAGTATTTCTGCAGAGCGGCCGTCGATGAAGACCCTGTTTGTGGACAGCTACAGTG
 AGATGCTTTTCTTCTGCAGTCACTGTTTATGCTGGCCACCGTGGTGTACTTCAGCC

ACCTCAAGGAGTATGTGGCTTCCATGGTATTCTCCCTGGCCTTGGGCTGGACCAACATGC
TCTACTACACCCGCGTTTCCAGCAGATGGGCATCTATGCCGTCATGATAGAGAAGATGA
TCCTGAGAGACCTGTGCCGTTTCATGTTTGTCTACATCGTCTTCTTGTTCGGGTTTTCCA
CAGCGGTGGTGACGCTGATTGAAGACGGGAAGAATGACTCCCTGCCGCTCTGAGTCCACGT
CGCACAGGTGGCGGGGCTGCCCTGCAGGCCCCCGATAGCTCCTACAACAGCCTGTACT
CCACCTGCCCTGGAGCTGTTCAAGTTCACCATCGGCATGGGCGACCTGGAGTTCACCTGAGA
ACTATGACTTCAAGGCTGTCTTCATCATCTCTGCTGCTGGCCTATGTAATTCTCACCTACA
TCCTCCTGCTCAACATGCTCATCGCCCTCATGGGTGAGACTGTCAACAAGATCGCACAGG
AGAGCAAGAACATCTGGAAGCTGCAGAGAGCCATCACCATCCTGGACACGGAGAAGAGCT
TCCTTAAAGTGCATGAGGAAGGCCTTCCGCTCAGGCAAGCTGCTGCAGGTGGGGTACACAC
CTGATGGCAAGGACGACTACCGGTGGTGCTTCAGGGTGGACGAGGTGAAGTGGACCACCT
GGAACACCAACGTGGGCATCATCAACGAAGACCCGGGCAACTGTGAGGGCGTCAAGCGCA
CCCTGAGCTTCTCCCTGCGGTCAAGCAGAGTTTCAGGCAGACACTGGAAGAACTTTGCC
TGGTCCCCCTTTAAGAGAGGCAAGTGTCTGAGATAGGCAGTCTGCTCAGCCCGAGGAAG
TTTATCTGCGACAGTTTTTCAGGGTCTCTGAAGCCAGAGGACGCTGAGGTCTTCAAGAGTC
CTGCCGCTTCCGGGAGAAAGTGAGGACGTACGCGAGACAGCACTGTCAACACTGGGCCTT
AGGAGACCCCGTTGCCACGGGGGCTGCTGAGGGAACACCAAGTGTCTGTGAGCAGCCTG
GCCTGGTCTGTGCTGCCAGCATGTTCCCAAATCTGTGCTGGACAAGCTGTGGGAAGCG
TTCTTGAAGCATGGGAGTGATGTACATCCAACCGTCACTGTCCCCAAGTGAATCTCCT
AACAGACTTTTCAAGTTTTTACTCACTTTACTAAACAGTKTGGATGGTCAGTCTCTACTGG
GACATGTTAGGCCCTTGTCTTTGATTTTATTCTTTTTTTGAGACAGAATTTCACTC
TTCTCACCCAGGCTGGAATGCAGTGGCACAATTTGGGCTCCCTGCAACCTCCGCCTCCTG
GATTCCAGCAATTTCTCCTGCCCTCGGCTTCCCAAGTAGCTGGGATTACAGGCACGTGCCAC
CATGTCTGGCTAATTTTTTGATTTTTTTAATAGATATGGGGTTTCGCCATGTTGGCCAG
GCTGGTCTCGAACTCCTGACCTCAGGTGATCCGCCACCTCGGCCTCCCAAAGTCTGGG
ATTACAGGTGTGAGCCTCCACACCTGGCTGTTTTCTTTGATTTTATTCTTTTTTTTTTT
TCTGTGAGACAGAGTTTCACTCTGTGTTGCCAGGCTGGAGTGCAGTGGTGTGATCTTGGC
TCACTGCAACTTCTGCCTCCCGGTTCAAGCGATTCTTCTGCTTCAGTCTCCCAAGTAGC
TTGGATTACAGGTGAGCACTACCACGCCCGGCTAATTTTTTGATTTTTTAATAGACGGG
GTTTCACCATGTTGGCCAGGCTGGTCTCGAACTCTTGACCTCAGGTGATCTGCCCCCTT
GGCCTCCCAAAGTGTGGGATTACAGGTGTGAGCCGCTGCGCTCGGCCTTCTTTGATTTT
ATATTATTAGGAGCAAAAGTAAATGAAGCCAGGAAAACACCTTTGGGAACAACTCTTC
CTTTGATGGAAAATGCAGAGGCCCTTCTCTCTGTGCGGTGCTTGTCTCTTACCTGCC
CGGGTGGTTTGGGGTGTGGTGTCTTCTCCCTGGAGAAGATGGGGGAGGCTGTCCCACT
CCCAGCTCTGGCAGAATCAAGCTGTTGCAGCAGTGCCTTCTTCATCCTTCTTACGATCA
ATCACAGTCTCCAGAAGATCAGCTCAATTGCTGTGCAGGTTAAACTACAGAACCACATC
CCAAAGGTACCTGGTAAGAATGTTTGAAGATCTTCCATTTCTAGGAACCCAGTCTGCTG
TTCTCCGCAATGGCACATGCTTCCACTCCATCCATACTGGCATCCTCAAATAAACAGATA
TGATACAATAAAAAAAAAAAAAAAAAAAAAARRGCGGCCGCTGAATTCTAGACCTGCCCGG
GCG

SEQ ID NO:2

MKXWSSTDLGAAADPLQKDTCPDPLDGPNSRPPPAKPKQLSTAKSRTRLFGKGDSEAF

VDCPHEEGELDCPTTITVSPVITIQRPDGPPTGARLLSQDSVAASTEKTLRLYDRRSIFE
 AVAQNQCQDLESLLLFLQKSKKHLTDNEFKDPETGKTCLLKAMLNLDGQNTTIPLLLEI
 ARQTDLSKELVNASYTDSYKQGTALHIAIERRNMLVTLLEVENGADVQAAAHGDFFKKT
 5 KGRPGFYFGEPLSLAACTNQLGIVKFLQNSWQTADISARDSVGNVTVLHALVEADNTA
 DNTKFVTSMYNEILILGAKLHPTLKLEELTNKKGMTPLALAAGTGKIGVLAYILQREIQE
 PECRHLRSKFTWAYGPVHSSLYDLSCIDTCEKNSVLEVIAYSSSETPNRHDMLLVEPLN
 RLLQDKWDRFVKRIFYFNFLVYCLYMIIFTMAAYRPPVDGLPPFKMEKTGDYFRVTGEIL
 10 SVLGGVYFFFRGIQYFLQRRPSMKTFLVDSYSEMLFPLQSLFMLATVVLYFSLKEYVAS
 MVFSLALGWTNMLYYTRGFQMGYAVMIEKMILRDLCRFMFVYIVFLFGFSTAVVTLIE
 DGKNDSLPSESTSHRWGPACRPPDSSYNSLYSTCLELFKFTIGMGDLEFTENYDFKAVF
 IILLLAYVILTYILLNMLIALMGETVNKIAQESKNIWKLQRAITILDTEKSLKCMRKA
 15 FRSGKLLQVGYTPDGKDDYRWCFRVDENVWTTWNTNVGIINEDPGNCEGVKRTLFSFLRS
 SRVSGRHWNFALVPLLRASARDRQSAQPEEVYLRQFSGSLKPDAEVFKSPAASGEK

SEQ ID NO:3

CAGCGTCGGGTGCAGTTTGGCCGGAGGTTGCAGTGAGCAGAGATTGCCCCATTGCACTCT
 AGTCTGGGCGACAGGGTGAGACACACACACACACACACACACACACACACACACAC
 25 ACACAAGCCTAAACATTCRAGGCCAGGATGCTTGACAGATGTTGATTCATAAAATGACA
 AAAAGCACAAAATCCAAAATCTCGTATAAGCTCAGTGGCTGTGGCAGCGAGGTTGAAGAG
 CAAAGGCAGGCCGGGCACCTGGCTGATGATGTGTGGACCCGTTGCACAGCAGGGCCCCGC
 AGTGGCGGTGTGGGTGTGGGTGTGGGTGGGCCAGTYTCTGCCGCTCACCTATTCCAGGGA
 CACAGTCTGCTTGGCTCTTCTGGACTGAGCCATCCTCATCACCGAGATCCTCCCTGAATT
 30 CAGCCCACGACAGCCACCCCGCCGTTTCTTGTCTGTGTGGGGAGGGAGGCAGCGCG
 GTGGTTATCAACCTCACCTGCAGAGGAGGCACCTGAGGCCACAGAGCAGGAGGGATGG
 GTCTAACCCAGAACCACAGATGGCTCTGAGCCGGGGGCTGTCCACCCTCCAGGCCGAC
 GTCAGTGGCCGACAGGACTGCCTGGGCCCTGCTAGGCCTGCTCACCTCTGAGGCCTCTGGG
 35 GTGAGAGGTTAGTCTCGAAACACTTCAGTTCTAGGGGGCTGGGGGCAGCAGCAAGTTG
 GAGTTTTGGGGTACCCTGCTTCACAGGCCCTTGGCAAGGAGGGCAGGTGGGGTCTAAGG
 ACAAGCAGTCTTACTTTGGGAGTCAACCCCGCGTGGTGGCTGCTGCAGGTTGCACACT
 GGGCCACAGAGGATCCAGCAAGGATGAAGAAATGGAGCAGCACAGACTTGGGGGCAGCTG
 CGGACCCACTCCAAAAGGACACCTGCCAGACCCCTGGATGGAGACCCTAACTCCAGGC
 40 CACCTCCAGCCAAGCCCAGCTCTCCACGGCCAAGAGCGCACCCGGCTCTTTGGGAAGG
 GTGACTCGGAGGAGGCTTTCCCGGTGGATTGCCCTCACAGGAAGGTGAGCTGGACTCCT
 GCCCGACCATCACAGTCAGCCCTGTTATCACCATCCAGAGGCCAGGAGACGGCCCCACCG
 GTGCCAGGCTGCTGTCCAGGACTCTGTGCGCGCCAGCACCGAGAAGACCTCAGGCTCT
 45 ATGATCGCAGGAGTATCTTTGAAGCCGTTGCTCAGAATAACTGCCAGGATCTGGAGAGCC
 TGCTGCTCTTCTGCAGAAGAGCAAGAAGCACYTCACAGACAACGAGTTCAAAGACCCTG
 AGACAGGGAAAGACCTGTCTGCTGAAAGCCATGCTCAACCTGCACGACGGACAGAACACCA
 CCATCCCCCTGCTCCTGGAGATCGCGGGCAAACGGACAGCCTGAAGGAGCTTGTC AACG
 50 CCRGCTACACGGACAGSTACTACAAGGGCCAGACAGCACTGCACATCGCCATCGAGAGAC
 GCAACATGGCCCTGGTGACCTCTCTGGTGAGAACGGAGCAGAGCTCCAGGCTGCGGCC
 ATGGGGACTTCTTTAAGAAAACCAAAGGGCGGCTGGATTCTACTTCGGTGAAGTCCCC

5
10
15
20
25
30
35
40
45
50
55

TGTCCTGGCCGCGTGACCAACCAGCTGGGCATCGTGAAGTTCTGCTGCAGAACTCCT
GGCAGACGGCCGACATCAGCGCCAGGACTCGGTGGGCAACACGGTGTGCACGCCCTGG
TGGAGGTGGCCGACAACACGGCCGACAACCAAGTTTGTGACGAGCATGTACAATGAGA
TTCTGATCCTGGGGGCCAACTGCACCCGACGCTGAAGCTGGAGGAGCTCACCAACAAGA
AGGGAATGACGCGCTGGCTCTGGCAGCTGGGACCGGGAAGATCGGGGTCTTGGCCTATA
TTCTCCAGCGGGAGATCCAGGAGCCGAGTGCAGGCACCTGTCCAGGAAGTTCACCGAGT
GGGCCTACGGGCCCGTGCACTCCTCGTGTACGACCTGTCTGCATCGACACCTGCGAGA
AGAACTCGGTGCTGGAGGTGATCGCCTACAGCAGCAGCGAGACCCCTAATCGCCACGACA
TGCTCTTGGTGGAGCCGCTGAACCGACTCCTGCAGGACAAGTGGGACAGATTTCGTC AAGC
GCATCTTCTACTTCAACTTCCTGGTCTACTGCCTGTACATGATCATCTTACCATGGCTG
CCTACTACAGGCCCGTGGATGGCTTGCTCCCTTTAAGATGGAAAAAAGTGGAGACTATT
TCCGAGTTACTGGAGAGATCCTGTCTGTGTTAGGAGGAGTCTACTTCTTTTCCGAGGGA
TTCAGTATTTCCTGCAGAGGCGGCCGTCGATGAAGACCCCTGTTTGTGGACAGCTACAGTG
AGATGCTTTTCTTCTGCAGTCACTGTTCATGCTGGCCACCGTGGTGTGTACTTCAGCC
ACCTCAAGGAGTATGTGGCTTCCATGGTATTCTCCCTGGCCTTGGGCTGGACCAACATGC
TCTACTACACCCCGGTTTCCAGCAGATGGGCATCTATGCCGTCATGATAGAGAAGATGA
TCCTGAGAGACCTGTGCCGTTTTCATGTTTGTCTACATCGTCTTCTTGTTCGGGTTTCCA
CAGCGGTGGTGACGCTGATTGAAGACGGGAAGAATGACTCCCTGCCGTCTGAGTCCACGT
CGCACAGGTGGCGGGGCTGCCTGCAGGCCCCCGATAGCTCCTACAACAGCCTGTACT
CCACCTGCCTGGAGCTGTTCAAGTTCACCATCGGCATGGGCGACCTGGAGTTCAGTGAGA
ACTATGACTTCAAGGCTGTCTTCATCATCCTGCTGTGGCTATGTAATTCTCACCTACA
TCCTCCTGCTCAACATGCTCATCGCCCTCATGGGTGAGACTGTCAACAAGATCGCACAGG
AGAGCAAGAACATCTGGAAGCTGCAGAGAGCCATCACCATCCTGGACACGAGAGAAGCT
TCCTTAAGTGCATGAGGAAGGCCCTCCGCTCAGGCAAGCTGCTGCAGGTGGGTACACAC
CTGATGGCAAGGACGACTACCGGTGGTGTCTCAGGGTGGACGAGGTGAACTGGACCACCT
GGAACACCAACGTTGGGCATCATCAACGAAGACCCGGGCAACTGTGAGGGCGTCAAGCGCA
CCCTGAGCTTCTCCTGCGGTCAAGCAGAGTTTCAGGCAGACACTGGAAGAACTTTGCC
TGGTCCCCCTTTAAGAGAGGCAAGTGTCTCGAGATAGGCAGTCTGCTCAGCCCGAGGAAG
TTTATCTGCGACAGTTTTCAGGGTCTCTGAAGCCAGAGGACGCTGAGGTCTTCAAGAGTC
CTGCCGCTTCCGGGGAGAAGTGAGGACGTCACGCAGACAGCACTGTCAACACTGGGCCTT
AGGAGACCCCGTTGCCACGGGGGCTGCTGAGGGAACACAGTGTCTGTGACGAGCCTG
GCCTGGTCTGTGCTGCCAGCATGTTCCCAATCTGTGCTGGACAAGCTGTGGGAAGCG
TTCTTGGAAGCATGGGGAGTGATGTACATCCAACCGTCACTGTCCCAAGTGAATCTCCT
AACAGACTTTCAGGTTTTTACTCACTTTACTAAACAGTKTGATGGTCAGTCTCTACTGG
GACATGTTAGGCCCTTGTCTTTGATTTTATTCTTTTTTTGAGACAGAATTTCACTC
TTCTCACCCAGGCTGGAATGCAGTGGCACAATTTGGCTCCCTGCAACCTCCGCCTCCTG
GATTCCAGCAATTCTCCTGCTCGGCTTCCCAAGTAGCTGGGATTACAGGCACGTGCCAC
CATGTCTGGCTAATTTTTTGTATTTTTTAATAGATATGGGGTTTCGCCATGTTGGCCAG
GCTGGTCTCGAACTCCTGACCTCAGGTGATCCGCCACCTCGGCCTCCCAAGTGTGGG
ATTACAGGTGTGAGCCTCCACACCTGGCTGTTTCTTTGATTTTATTCTTTTTTTTTTT
TCTGTGAGACAGAGTTTCACTCTTGTGCCCAGGCTGGAGTGCAGTGGTGTGATCTTGGC
TCACTGCAACTTCTGCCTCCCGGGTTCAAGCGATTCTTCTGCTTCAGTCTCCCAAGTAGC
TTGGATTACAGGTGAGCACTACCACGCCCGGCTAATTTTTGTATTTTAATARAGACGGG

GTTTCACCATGTTGGCCAGGCTGGTCTCGAACTCTTGACCTCAGGTGATCTGCCCCCCTT
 GGCCTCCCAAAGTCTGGGATTACAGGTGTGAGCCGCTGCGCTCGGCCTTCTTTGATTTT
 ATATTATTAGGAGCAAAAGTAAATGAAGCCAGGAAACACCTTTGGGAACAAACTCTTC
 CTTTGATGGAAAATGCAGAGGCCCTTCCTCTCTGTGCCGTGCTTGCTCCTTTACCTGCC
 CGGGTGGTTTGGGGGTGTTGGTGTTCCTCCCTGGAGAAGATGGGGGAGGCTGTCCCACT
 CCCAGCTCTGGCAGAATCAAGCTGTTGCAGCAGTGCCTTCTTCATCCTTCCTTACGATCA
 ATCAGAGTCTCCAGAAGATCAGCTCAATTGCTGTGCAGGTTAAACTACAGAACCACATC
 CCAAAGGTACCTGGTAAGAATGTTTGAAGATCTTCCATTCTAGGAACCCAGTCCTGC
 TTCTCCGCAATGGCACATGCTTCCACTCCATCCATACTGGCATCCTCAAATAAACAGATA
 TGTATACAATAAAAAAAAAAAAAAAAAARRGCGGCCGCTGAATTCTAGACCTGCCCGG
 GCG

SEQ ID NO:4

MKKWSSTD LGAAADPLQKDTCPDPLDGPNSRPPPAKPQLSTAKSRTRLFGKGDSEAFP
 VDCPHEEGELDSCPTITVSPVITIQRPGDGP TGARLLSQDSVAASTEKT LRLYDRRSIFE
 AVAQNNQCQDLESLLLFLQKSKHXDTNEFKDPETGKTCLLKAMLNLHDGQNTTIPLLLEI
 ARQTDLSLKELVNAXYTDXYKQGQTALHIAIERNNMALVTLLEVENGADVQAAAHGDFFKT
 KGRPGFYFGELPLSLAACTNQLGIVKFLQNSWQTADISARDSVGNVTLHALVEVADNTA
 DNTKFVTSMYNEILILGAKLHPTLKLEELTNKKGMTPLALAAGTGKIGVLAYILQREIQE
 PECRHLRSRKFTWAYGPFVHSSLYDLSCIDTCEKNSVLEVIAYSSSETPNRHDMLLVEPLN
 RLLQDKWDRFVKRIFYFNFLVYCLYMIIFTMAAYRPPVDGLPPFKMEKTGDYFRVTGEIL
 SVLGGVYFPFRGIQYFLQRRPSMKTFLVDSYSEMLFFLQSLFMLATVVLYFSLHKEYVAS
 MVFSLALGWTNMLYYTRGFQQMGIYAVMIEKMILRDLCRPFVYIVFLFGFSTAVVTLIE
 DGKNDSLPSESTSHRWGPACRPPDSSYNSLYSTCLELFKFTIGMGDLEPTENYDFKAVF
 IILLLAYVILTYIILLNMLIALMGETVNKIAQESKNIWKLQRAITILDTEKSFLKCMRKA
 FRSGKLLQVGYTPDGKDDYRWCFRVEVNWTTWNTNVGIINEDPGNCEGVKRTLFSFLRS
 SRVSGRHWKNFALVPLLREASARDRQSAQPEEVYLRQFSGSLKPDAEVFKSPAASGEK

SEQ ID NO:5

GAGCTTCTCCCTGCCGTCAAGCAGAGTTTCAGGCAGACACTGGAAGAACTTTGCCCTGGT
 CCCCCCTTTAAGAGAGGCAAGTNCTCGANATAGGCAGTCTGCTCAGCCCGAGGAAGTTTA
 TCTGCGACAGTTTTCAGGGTCTCTAAAGCCAGAGGACGCTGAGGTCTTCAAGAGTCCTGC
 CGCTTCCGGGGAGAAGTGAGGACGTCACGCAGACAGCACTGTCAACACTGGGCCTTAGGA
 GACCCCGTTGCCACGGGGGGCTGCTGAGGGAACACCAGTGCTTTTTCAGCAGCCTTGCCCT
 GGGTCTTTGCCCTGCCAGCATGTTCCCAAATCTGTGCTGGACAAGCTGTGGGGAAGCGTT
 CTTGGGAAGCATGGGGGAGTGATGTTACATCCAACCGTCACTGTCCCAAGTTGAATCTT
 CCTTAACAGATT

SEQ ID NO:6

SFSLRSSRVSGRHWKNFALVPLLREASXRQSAQPEEVYLRQFSGSLKPDAEVFKSPA
 ASGEK

SEQ ID NO:7, PVP-1 (polymorphic variant of VANILREP1, A2625G)

5 CAGCGTCGGGTGCAGTTTGGCCGGAGGTTGCAGTGAGCAGAGATTGCCCCATTGCACCTCT
 AGTCTGGGCGACAGGGTGAGACACACACACAGACACACACACACACACACACAC
 ACACAAGCCTAAACATTTCRAGGCCAGGATGCTTGACAGATGTTGATTCTATAAAATGACA
 AAAAGCACAAAATCCAAAATCTCGTATAAGCTCAGTGGCTGTGGCAGCGAGGTTGAAGAG
 10 CAAAGGCAGGCCGGGCACCTGGCTGATGATGTGTGGACCCGTTGCACAGCAGGGCCCCGC
 AGTGCAGGTGTGGGTGTGGGTGTGGGTGGGCCAGTYTCTGCCGCTCACCTATTCCAGGGA
 CACAGTCTGCTTGGCTCTTCTGGAAGTGCAGCCATCCTCATACCGAGATCCTCCCTGAATT
 CAGCCCACGACAGCCACCCCGCCGTTTTCTTGTCTGTGTGGGAGGGAGGCAGCGCG
 GTGGTTATCAACCTCACCTGCAGAGGAGGCACCTGAGGCCAGAGACGAGGAGGGATGG
 15 GTCTAACCCAGAACCACAGATGGCTCTGAGCCGGGGCCCTGTCCACCCTCCCAGGCCGAC
 GTCAGTGGCCGACGACTGCCTGGGCCCTGCTAGGCCTGCTCACCTCTGAGGCCTCTGGG
 GTGAGAGGTTCACTCCTGGAAACACTTCAGTTCTAGGGGGCTGGGGGCAGCAGCAAGTTG
 GAGTTTGGGGTACCCTGCTTCACAGGGCCCTTGGCAAGGAGGGCAGGTGGGGTCTAAGG
 20 ACAAGCAGTCTTACTTTGGGAGTCAACCCCGGCGTGGTGGCTGCTGCAGGTTGCACACT
 GGGCCACAGAGGATCCAGCAAGGATGAAGAAATGGAGCAGCACAGACTTGGGGGCAGCTG
 CGGACCCACTCCAAAAGGACACCTGCCAGACCCCTGGATGGAGACCCTAACTCCAGGC
 CACCTCCAGCCAAGCCCCAGCTCTCCAGGCCAAGAGCCGACCCCGCTCTTTGGGAAGG
 25 GTGACTCGGAGGAGGCTTTCCCGGTGGATTGCCCTCACAGGAAGGTGAGCTGGACTCCT
 GCCCCACCATCACAGTCAGCCCTGTTATCACCATCCAGAGGCCAGGAGACGGCCCCACCG
 GTGCCAGGCTGCTGTCCCAGGACTCTGTGCGCGCCAGCACCCAGGAAGACCCTCAGGCTCT
 ATGATCGCAGGAGTATCTTTGAAGCCGTTGCTCAGAATAACTGCCAGGATCTGGAGAGCC
 30 TGCTGCTCTTCCTGCAGAAGAGCAAGAAGCACCTCACAGACAACGAGTTCAAAGACCCTG
 AGACAGGGAAGACCTGTCTGCTGAAAGCCATGCTCAACCTGCACGACGGACAGAACACCA
 CCATCCCCCTGCTCCTGGAGATCGCGCGGCAACGGACAGCCTGAAGGAGCTTGTCAACG
 CCAGCTACACGGACAGCTACTACAAGGGCCAGACAGCACTGCACATCGCCATCGAGAGAC
 35 GCAACATGGCCCTGGTGACCTCCTGGTGGAGAACGGAGCAGACGTCAGGCTGCGGCCC
 ATGGGGACTTCTTTAAGAAAACCAAAGGGCGGCGCTGGATTCTACTTCGGTGAAGTGCCCC
 TGTCCCTGGCCGCGTGACCAACAGCTGGGCATCGTGAAGTTCTGCTGCAGAACTCCT
 GGCAGACGGCCGACATCAGCGCCAGGACTCGGTGGGCAACACGGTGTGCACGCCCTGG
 40 TGGAGGTGGCCGACAACACGGCCGACAACACGAAGTTGTGACGAGCATGTACAATGAGA
 TTCTGATCCTGGGGGCCAACTGCACCCGACGCTGAAGCTGGAGGAGCTCACCAACAAGA
 AGGGAAATGACCGCGCTGGCTCTGGCAGCTGGGACCGGGAAGATCGGGGTCTTGGCCTATA
 TTCTCCAGCGGGAGATCCAGGAGCCCGAGTGCAGGCACCTGTCCAGGAAGTTACCGAGT
 45 GGGCCTACGGGCCCGTGCACTCCTCGCTGTACGACCTGTCTGCATCGACACCTGCGAGA
 AGAACTCGGTGCTGGAGGTGATCGCCTACAGCAGCAGCGAGACCCCTAATCGCCACGACA
 TGCTCTTGGTGGAGCCGCTGAACCGACTCCTGCAGGACAAGTGGGACAGATTTCGTCAAGC
 GCATCTTCTACTTCAACTTCCTGGTCTACTGCCTGTACATGATCATCTTCACCATGGCTG
 50 CCTACTACAGGCCCGTGATGGCTTGCTCCCTTTAAGATGGAAAAAATGGGAGACTATT
 TCCGAGTTACTGGAGAGATCCTGTCTGTGTTAGGAGGAGTCTACTTCTTTTCCGAGGGA
 TTCAGTATTTCTGTCAGAGGCGGCCGTCGATGAAGACCCTGTTTGTGGACAGCTACAGTG
 AGATGCTTTTCTTTCTGCAGTCACTGTTTCATGCTGGCCACCGTGGTGTGACTTCAGCC

55

ACCTCAAGGAGTATGTGGCTTCCATGGTATTCTCCCTGGCCTTGGGCTGGACCAACATGC
TCTACTACACCCGCGGTTTCCAGCAGATGGGCATCTATGCCGTCATGATAGAGAAGATGA
5 TCCTGAGAGACCTGTGCCGTTTCATGTTGTCTACGTCGTCTTCTTGTTCGGGTTTCCCA
CAGCGGTGGTGACGCTGATTGAAGACGGAAGAATGACTCCCTGCCGTCTGAGTCCACGT
CGCACAGGTGGCGGGGGCCTGCCTGCAGGCCCCCGATAGCTCCTACAACAGCCTGTACT
CCACCTGCCTGGAGCTGTTCAAGTTCACCATCGGCATGGGCGACCTGGAGTTCACTGAGA
10 ACTATGACTTCAAGGCTGTCTTCATCATCCTGCTGCTGGCCTATGTAATTCTCACCTACA
TCCTCCTGCTCAACATGCTCATCGCCCTCATGGGTGAGACTGTCAACAAGATCGCACAGG
AGAGCAAGAACATCTGGAAGCTGCAGAGAGCCATCACCATCCTGGACACGGAGAAGAGCT
TCCTTAAGTGCATGAGGAAGGCCTTCCGCTCAGGCAAGCTGCTGCAGGTGGGGTACACAC
15 CTGATGGCAAGGACGACTACCGGTGGTGCTTCAGGGTGGACGAGGTGAAGTGGACCACCT
GGAACACCAACGTGGGCATCATCAACGAAGACCCGGGCAACTGTGAGGGCGTCAAGCGCA
CCCTGAGCTTCTCCCTGCGGTCAAGCAGAGTTTCAGGCAGACACTGGAAGAATTGCCC
TGGTCCCCCTTTTAAGAGAGGCAAGTGCTCGAGATAGGCAGTCTGCTCAGCCGAGGAAG
20 TTTATCTGCGACAGTTTTCAGGGTCTCTGAAGCCAGAGGACGCTGAGGTCTTCAAGAGTC
CTGCCGCTTCCGGGGAGAAGTGAGGACGTCACGCAGACAGCACTGTCAACACTGGGCCTT
AGGAGACCCCGTTGCCACGGGGGGCTGCTGAGGGAACACCAGTGCTCTGTGACAGCCTG
25 GCCTGGTCTGTGCCTGCCCA

SEQ ID NO:8, (PVP-1 protein, I585V)

MKKWSSTDLGAAADPLQKDTCPDPLDGDPNRPPPAKPQLSTAKSRTLFGKGDSEEAFF
30 VDCPHEEGELDSCPTITVSPVITIQRPDGP TGARLLSQDSVAASTEKTLRLYDRSIFE
AVAQNNCQDLESLLLFLQSKKHLTDNEFKDPETGKTCLLKAMNLHDCQNTTIPLLLEI
ARQTDLSKELVNASYTDSYKQGTALHIAIERRMALVTLLVENGADVQAAAHGDFFKKT
KGRPGFYFGEPLPLSLA ACTNQLGIVKFLQNSWQTADISARDSVGNTVLHALVEVADNTA
35 DNTKFVTSMYNEILILGAKLHPTLKLEELTNKGMTPLALAAGTGKIGVLAYILQREIQE
PECRHLSRKFTWAYGPFVHSSLYDLSCIDTCEKNSVLEVIAYSSSETPNRHDMMLLVEPLN
RLQDKWDRFVKRIFYFNFLVYCLYMIIFTMAAYRPPVDGLPPFKMEKTGDYFRVTGEIL
SVLGGVYFFFRGIQYFLQRRPSMKTLFVDSYSEMLFFLQSLFMLATVVLYFSLKEYVAS
40 MVFSLALGWTNMLYYTRGFQQMGIYAVMIEKMILRDLCRFMFVYVFLFGFSTAVVTLIE
DGKNDLPSSESTSHRWGPACRPPDSSSYNSLYSTCLELFKFTIGMGDLEFTENYDFKAVF
IILLAYVILTYIILLNMLIALMGETVNKIAQESKNIWKLQRAITILDTEKSFLKCMRKA
45 FRSGKLLQVGYPDGKDDYRWCFRVDEVNWTWNTNVGIINEDPGNCEGVKRTLSFSLRS
SRVSGRHWKNFALVPLLEASARDRQSAQPEEVYLRQFSGSLKPEDAEVFKSPAASGEK

SEQUENCE LISTING

5 <110> SmithKline Beecham plc

<120> Novel Compounds

10 <130> GP30075

<160> 8

15 <170> FastSEQ for Windows Version 3.0

<210> 1

20 <211> 4803

<212> DNA

<213> Homo sapiens

25 <400> 1

cagcgtcggg tgcagtttgg ccggagggtg cagtgcagcag agattgcccc attgcactct 60

agtctgggcg acaggggtgag acacacacac acagacacac acacacacac acacacacac 120

acacaagcct aaacattcra ggccaggatg cttgacagat gttgattcat aaaaatgaca 180

30 aaaagcacia aatccaaaat ctctgtataag ctctagtggt gtggcagcga ggttgaagag 240

caaaggcagg ccggggcacct ggctgatgat gtgtggaccc gttgcacagc agggccccgc 300

agtgcggtgt ggggtgtgggt gtgggtgggc cagtytctgc cgtccacctt attccaggga 360

cacagtctgc ttggctcttc tggactgagc catcctcatc accgagatcc tccctgaatt 420

35 cagcccacga cagccacccc ggccgttttc cttgttctgt gtggggaggg aggcagcgcg 480

gtggttatea acctcacctt gcagaggagg cacctgaggc ccagagacga ggagggatgg 540

gtctaacca gaaccacaga tggctctgag ccgggggcct gtccaccttc ccaggccgac 600

gtcagtggcc gcaggactgc ctgggacctg ctaggacctgc tcacctctga ggacctctggg 660

40 gtgagagggt cagtcctgga aacacttcag ttctaggggg ctgggggcag cagcaagttg 720

gagttttggg gtaccttgct tcacaggguc cttggcaagg agggcagggt gggctctaagg 780

acaagcagtc cttactttgg gagtcaacct cggcgtggtg gctgctgcag gttgcacact 840

ggggcacaga ggaaccagca aggatgaaga aatggagcag cacagacttg ggggcagctg 900

45 cggacccact ccaaaaggac acctgcccag accccctgga tggagacctt aactccaggc 960

cacctccagc caagccccag ctctccacgg ccaagagccg caccggcttc ttggggaagg 1020

gtgactcgga ggaggcttcc ccggtggatt gccctcacga ggaaggtag ctggactcct 1080

gcccagccat cacagtcagc cctgttatea ccattccagag gccaggagac ggccccaccg 1140

50 gtgccaggct gctgtcccag gactctgtcg ccgccagcac cgagaagacc ctccaggctct 1200

atgatcgag gagtatcttt gaagccgttg ctcaagaata ctgccaggat ctggagagcc 1260

tgtgtctctt cctgcagaag agcaagaagc acctcacaga caacgagttc aaagaccctg 1320

55

	agacaggga gacctgtctg ctgaaagcca tgctcaacct gcacgacgga cagaacacca	1380
	ccatccccct gctcctggag atcgcgcggc aaacggacag cctgaaggag cttgtcaacg	1440
5	ccagctacac ggacagctac tacaagggcc agacagcact gcacatcgcc atcgagagac	1500
	gcaacatggc cctggtgacc ctctggtgg agaacggagc agacgtccag gctgcggccc	1560
	atggggactt ctttaagaaa accaaagggc ggcttgatt ctacttcggt gaactgcccc	1620
	tgccccggc cgcgtgcacc aaccagctgg gcacgtgaa gttcctgctg cagaactcct	1680
10	ggcagacggc cgacatcagc gccagggact cggtagggcaa cagggtgctg cagccctgg	1740
	tggagggtggc cgacaacacg gccgacaaca cgaagtttgt gacgagcatg tacaatgaga	1800
	ttctgatcct gggggccaaa ctgcaccgca cgtgaagct ggaggagctc accaacaaga	1860
	agggaaatgac gccgctggct ctggcagctg ggaccgggaa gatcggggtc ttggcctata	1920
15	ttctccagcg ggagatccag gagcccgagt gcaggcacct gtccaggaaq ttaccgagt	1980
	gggcctacgg gcccgctgcac tcctcgctgt acgacctgtc ctgcatcgac acctgcgaga	2040
	agaactcggg gctggagggt atgcctaca gcagcagcga gaccctaact cgccacgaca	2100
	tgctcttggt ggagccgctg aaccgactcc tgcaggacaa gtgggacaga ttctcaagc	2160
20	gcactctcta ctcaacttc ctggtctact gcctgtacat gatcatcttc acctggctg	2220
	cctactacag gcccgctggat ggcttgcttc ctttaagat ggaaaaaact ggagactatt	2280
	tccgagttac tggagagatc ctgtctgtgt taggaggagt ctacttcttt ttccgaggga	2340
	ttcagtatth cctgcagagg cggccgctga tgaagaccct gtttgtagac agctacagt	2400
25	agatgctttt ctttctgcag tcactgttca tgctggccac cgtggtgctg tacttcagcc	2460
	acctcaagga gtatgtggct tccatggtat tctccctggc cttgggctgg accaactgc	2520
	tctactacac ccgcggttcc cagcagatgg gcacttatgc cgtcatgata gagaagatga	2580
	tcctgagaga cctgtgcctg ttcatgttg tctacatcgt cttcttgctc ggggtttcca	2640
	cagcgggtgt gacgctgatt gaagacggga agaatactc cctgccgtct gagtccact	2700
30	cgacaggtg gcgggggctt gcctgcaggg ccccgatag ctctacaac agcctgtact	2760
	ccacctgcct ggagctgttc aagttacca tggcatggg cgaacctggg ttactgaga	2820
	actatgactt caaggctgtc ttcatcctc tgcgtctggc ctatgtaatt ctacactaca	2880
35	tcctctgct caacatgctc atgcctcca tgggtgagac tgtcaacaag atcgacagg	2940
	agagcaagaa catctggaag ctgcagagag ccatcaccat cctggacacg gagaagagct	3000
	tccttaagtg catgaggaag gccttccgt caggcaagct cctgcaggtg gggtagacac	3060
	ctgatggcaa ggacgactac cgggtggtgt tcagggtgga cgaagtgaa tggaccacct	3120
40	ggaacaccaa cgtgggcatc atcaacgaag acccgggcaa ctgtgagggc gtcaagcgca	3180
	ccctgagctt ctccctgcgg tcaagcagag tttaggcag aactggaag aactttgccc	3240
	tgggtccccct tttaagagag gcaagtgtc gagataggca gtctgctcag cccgaggaag	3300
	tttatctgag acagttttca gggctctga agccagagga cgtgaggtc ttcaagagtc	3360
	ctgccgtctc cggggagaag tgaggagctc acgcagacag cactgtcaac actgggcctt	3420
45	aggagacccc gtgcccacgg ggggtctctg agggaaacac agtgctctgt cagcagcctg	3480
	gcctggtctg tgcctgccc gcctgttccc aaatctgtgc tggacaagct gtgggaagcg	3540
	ttcttggaag catggggagt gatgtacac caaccgtcac tgtcccaag tgaatctcct	3600
50	aacagacttt cagggtttta ctactttac taaacagtkt ggatggctag tctctactgg	3660
	gacatgttag gccctgtttt tctttgatit tattcttttt tttagacag aattttactc	3720
	ttctcaccca ggctggaatg cagtggcaca attttggctc cctgcaacct ccgcctcctg	3780
	gattccagca atttctctgc ctggcttcc caagtagctg ggattacagg cactgcccac	3840

55

catgtctggc taattttttg tattttttta atagatatgg ggtttcgcca tgttggccag 3900
 gctgggtctcg aactcctgac ctcagggtgat ccgcccacct eggcctccca aagtgtgagg 3960
 5 attacagggtg tgagcctcca cacctggctg ttttctttga ttttattctt tttttttttt 4020
 tctgtgagac agagtttcac tcttgttgcc cagggtggag tgcagtgggtg tgatcttgcc 4080
 tcaactgaac tcttgcctcc cgggttcaag cgattcttct gcttcagtct cccaagtagc 4140
 ttggattaca ggtgagcact accacgcccg gctaattttt gtatttttaa taragacggg 4200
 10 gtttcaccat gttggccagg ctggtctcga actcttgacc tcaggtgacg tgcccgcctt 4260
 ggctcccaa agtgctggga ttacagggtg gagccgctgc gctcggcctt ctttgatttt 4320
 atattattag gagcaaaagt aaatgaagcc caggaaaaca ctttgggaa caaactcttc 4380
 ctttgatgga aaatgcagag gcccttcctc tctgtgccgt gcttgcctct cttacctgcc 4440
 15 cgggtggttt gggggtgttg gtgtttctc cctggagaag atgggggagg ctgtccact 4500
 cccagctctg gcagaatcaa gctgtgcag cagtgccttc ttcactcttc cttacgatca 4560
 atcacagtct ccagaagatc agctcaattg ctgtgcaggt taaaactaca gaaccacatc 4620
 ccaaaggtag ctggttaagaa tgtttgaaag atcttcatt tctaggaacc ccagtcctgc 4680
 20 ttctcgcgaa tggcacatgc ttccactcca tccatactgg catcctcaaa taaacagata 4740
 tgtatacaat aaaaaaaaaa aaaaaaaaaa rrgcggccgc tgaattctag acctgcccgg 4800
 gcg 4803

25 <210> 2
 <211> 839
 <212> PRT
 <213> Homo sapiens

30 <400> 2
 Met Lys Lys Trp Ser Ser Thr Asp Leu Gly Ala Ala Ala Asp Pro Leu
 1 5 10 15
 35 Gln Lys Asp Thr Cys Pro Asp Pro Leu Asp Gly Asp Pro Asn Ser Arg
 20 25 30
 Pro Pro Pro Ala Lys Pro Gln Leu Ser Thr Ala Lys Ser Arg Thr Arg
 35 40 45
 40 Leu Phe Gly Lys Gly Asp Ser Glu Glu Ala Phe Pro Val Asp Cys Pro
 50 55 60
 His Glu Glu Gly Glu Leu Asp Ser Cys Pro Thr Ile Thr Val Ser Pro
 65 70 75 80
 Val Ile Thr Ile Gln Arg Pro Gly Asp Gly Pro Thr Gly Ala Arg Leu
 45 85 90 95
 Leu Ser Gln Asp Ser Val Ala Ala Ser Thr Glu Lys Thr Leu Arg Leu
 100 105 110
 Tyr Asp Arg Arg Ser Ile Phe Glu Ala Val Ala Gln Asn Asn Cys Gln
 50 115 120 125
 Asp Leu Glu Ser Leu Leu Phe Leu Gln Lys Ser Lys Lys His Leu
 130 135 140

55

Thr Asp Asn Glu Phe Lys Asp Pro Glu Thr Gly Lys Thr Cys Leu Leu
 145 150 155 160
 Lys Ala Met Leu Asn Leu His Asp Gly Gln Asn Thr Thr Ile Pro Leu
 5 165 170 175
 Leu Leu Glu Ile Ala Arg Gln Thr Asp Ser Leu Lys Glu Leu Val Asn
 180 185 190
 Ala Ser Tyr Thr Asp Ser Tyr Tyr Lys Gly Gln Thr Ala Leu His Ile
 10 195 200 205
 Ala Ile Glu Arg Arg Asn Met Ala Leu Val Thr Leu Leu Val Glu Asn
 210 215 220
 Gly Ala Asp Val Gln Ala Ala Ala His Gly Asp Phe Phe Lys Lys Thr
 15 225 230 235 240
 Lys Gly Arg Pro Gly Phe Tyr Phe Gly Glu Leu Pro Leu Ser Leu Ala
 245 250 255
 Ala Cys Thr Asn Gln Leu Gly Ile Val Lys Phe Leu Leu Gln Asn Ser
 20 260 265 270
 Trp Gln Thr Ala Asp Ile Ser Ala Arg Asp Ser Val Gly Asn Thr Val
 275 280 285
 Leu His Ala Leu Val Glu Val Ala Asp Asn Thr Ala Asp Asn Thr Lys
 25 290 295 300
 Phe Val Thr Ser Met Tyr Asn Glu Ile Leu Ile Leu Gly Ala Lys Leu
 305 310 315 320
 His Pro Thr Leu Lys Leu Glu Glu Leu Thr Asn Lys Lys Gly Met Thr
 30 325 330 335
 Pro Leu Ala Leu Ala Ala Gly Thr Gly Lys Ile Gly Val Leu Ala Tyr
 340 345 350
 Ile Leu Gln Arg Glu Ile Gln Glu Pro Glu Cys Arg His Leu Ser Arg
 35 355 360 365
 Lys Phe Thr Glu Trp Ala Tyr Gly Pro Val His Ser Ser Leu Tyr Asp
 370 375 380
 Leu Ser Cys Ile Asp Thr Cys Glu Lys Asn Ser Val Leu Glu Val Ile
 40 385 390 395 400
 Ala Tyr Ser Ser Ser Glu Thr Pro Asn Arg His Asp Met Leu Leu Val
 405 410 415
 Glu Pro Leu Asn Arg Leu Leu Gln Asp Lys Trp Asp Arg Phe Val Lys
 45 420 425 430
 Arg Ile Phe Tyr Phe Asn Phe Leu Val Tyr Cys Leu Tyr Met Ile Ile
 435 440 445
 Phe Thr Met Ala Ala Tyr Tyr Arg Pro Val Asp Gly Leu Pro Pro Phe
 50 450 455 460
 Lys Met Glu Lys Thr Gly Asp Tyr Phe Arg Val Thr Gly Glu Ile Leu
 465 470 475 480

55

Ser Val Leu Gly Gly Val Tyr Phe Phe Phe Arg Gly Ile Gln Tyr Phe
 485 490 495
 5 Leu Gln Arg Arg Pro Ser Met Lys Thr Leu Phe Val Asp Ser Tyr Ser
 500 505 510
 Glu Met Leu Phe Phe Leu Gln Ser Leu Phe Met Leu Ala Thr Val Val
 515 520 525
 10 Leu Tyr Phe Ser His Leu Lys Glu Tyr Val Ala Ser Met Val Phe Ser
 530 535 540
 Leu Ala Leu Gly Trp Thr Asn Met Leu Tyr Tyr Thr Arg Gly Phe Gln
 545 550 555 560
 15 Gln Met Gly Ile Tyr Ala Val Met Ile Glu Lys Met Ile Leu Arg Asp
 565 570 575
 Leu Cys Arg Phe Met Phe Val Tyr Ile Val Phe Leu Phe Gly Phe Ser
 580 585 590
 20 Thr Ala Val Val Thr Leu Ile Glu Asp Gly Lys Asn Asp Ser Leu Pro
 595 600 605
 Ser Glu Ser Thr Ser His Arg Trp Arg Gly Pro Ala Cys Arg Pro Pro
 610 615 620
 25 Asp Ser Ser Tyr Asn Ser Leu Tyr Ser Thr Cys Leu Glu Leu Phe Lys
 625 630 635 640
 Phe Thr Ile Gly Met Gly Asp Leu Glu Phe Thr Glu Asn Tyr Asp Phe
 645 650 655
 30 Lys Ala Val Phe Ile Ile Leu Leu Leu Ala Tyr Val Ile Leu Thr Tyr
 660 665 670
 Ile Leu Leu Leu Asn Met Leu Ile Ala Leu Met Gly Glu Thr Val Asn
 675 680 685
 35 Lys Ile Ala Gln Glu Ser Lys Asn Ile Trp Lys Leu Gln Arg Ala Ile
 690 695 700
 Thr Ile Leu Asp Thr Glu Lys Ser Phe Leu Lys Cys Met Arg Lys Ala
 705 710 715 720
 40 Phe Arg Ser Gly Lys Leu Leu Gln Val Gly Tyr Thr Pro Asp Gly Lys
 725 730 735
 Asp Asp Tyr Arg Trp Cys Phe Arg Val Asp Glu Val Asn Trp Thr Thr
 740 745 750
 45 Trp Asn Thr Asn Val Gly Ile Ile Asn Glu Asp Pro Gly Asn Cys Glu
 755 760 765
 Gly Val Lys Arg Thr Leu Ser Phe Ser Leu Arg Ser Ser Arg Val Ser
 770 775 780
 50 Gly Arg His Trp Lys Asn Phe Ala Leu Val Pro Leu Leu Arg Glu Ala
 785 790 795 800
 Ser Ala Arg Asp Arg Gln Ser Ala Gln Pro Glu Glu Val Tyr Leu Arg
 805 810 815

55

EP 0 943 683 A1

Gln Phe Ser Gly Ser Leu Lys Pro Glu Asp Ala Glu Val Phe Lys Ser
 820 825 830
 Pro Ala Ala Ser Gly Glu Lys
 835

<210> 3
 <211> 4803
 <212> DNA
 <213> Homo sapiens

<400> 3

15	cagcgtcggg tgcagtttgg ccggagggtg cagtgcagcag agattgcccc attgcactct	60
	agtcctggggc acagggtgag acacacacac acagacacac acacacacac acacacacac	120
	acacaagcct aaacattcra ggccaggatg cttgacagat gttgattcat aaaaatgaca	180
	aaaagcacia aatccaaaat ctctataag ctgagtggtc gtggcagcga ggttgaagag	240
20	caaaggcagg ccggggcacct ggctgatgat gtgtggaccc gttgcacagc agggccccgc	300
	agtgcggtgt ggggtgtgggt gtgggtgggc cagtgtctgc cgctcacctc attccaggga	360
	cacagtcctgc ttggctcttc tggactgagc catcctcatc accgagatcc tccctgaatt	420
	cagcccacga cagccacccc ggccgttttc cttgttctgt gtggggaggg aggcagcgcg	480
25	gtggttatca acctcacctc gcagaggagg cacctgaggc ccagagacga ggagggatgg	540
	gtctaaccga gaaccacaga tggctctgag ccgggggcct gtccacccctc ccaggccgac	600
	gtcagtgccc gcaggactgc ctgggcccctg ctaggcctgc tcacctctga ggctcttggg	660
	gtgagaggtt cagtcttgga aacacttcag ttctaggggg ctgggggcag cagcaagttg	720
30	gagttttggg gtaccctgct tcacagggcc cttggcaagg agggcagggtg gggcttaagg	780
	acaagcagtc cttacttttg gagtcaaccc cggcgtggtg gctgctgcag gttgcacact	840
	gggcccacga ggatccagca aggatgaaga aatggagcag cacagacttg ggggcagctg	900
	cggaccactc ccaaaaggac acctgccagc acccctgga tggagacctc aactccaggc	960
35	cacctccagc caagccccag ctctccacgg ccaagagccg caccggctc tttgggaagg	1020
	gtgactcga gagggtttc ccggtggatt gccctcacga ggaaggtag ctggactcct	1080
	gcccagccat cacagtcagc cctgttatca ccatccagag gccaggagac ggccccaccg	1140
	gtgccagggt gctgtcccag gactctgtcg ccgccagcac cgagaagacc ctcaggctct	1200
40	atgatcgagc gattatcttt gaagccgttg ctcaagaata ctgccaggat ctggagagcc	1260
	tgctgtctct cctgcagaag agcaagaagc acytacaga caacgagttc aaagacctg	1320
	agacagggaa gacctgtctg ctgaaagcca tgcacaacct gcacgacgga cagaacacca	1380
	ccatccccct gctcctggag atcgcgcggc aaacggacag cctgaaggag cttgtcaacg	1440
45	ccrgctacac ggacagstac tacaagggcc agacagcact gcacatcgcc atcgagagac	1500
	gcaacatggc cctggtgacc ctccctgttg agaacggagc agacgtccag gctgcggccc	1560
	atggggactt ctttaagaaa accaaagggc ggccctggatt ctacttcggg gaactgcccc	1620
	tgtccctggc cgcgtgcacc aaccagctgg gcatcgtgaa gttcctgctg cagaactcct	1680
50	ggcagacggc cgacatcagc gccagggact cgggtgggcaa cccggtgctg cagcctctgg	1740
	tggagggtgg cgacaacacg gccgacaaca cgaagtttgt gacgagcatg tacaatgaga	1800
	ttctgatcct gggggccaaa ctgcaccgca cgtcgaagct ggaggagctc accaacaaga	1860

	agggaaatgac gccgctggct ctggcagctg ggaccgggaa gatcgggggtc ttggcctata	1920
	ttctccacgg ggagatccag gagcccgagt gcaggcacct gtccaggaaag ttcaccgagt	1980
5	gggcctacgg gcccggtcac tctctcgtgt acgacctgtc ctgcatcgac acctgcgaga	2040
	agaactcggg gctggagggt atcgccctaca gcagcagcga gacccctaata cgccacgaca	2100
	tgctcttggg ggagccgctg aaccgactcc tgcaggacaa gtgggacaga ttctgtaagc	2160
	gcactcttcta ctccaacttc ctgggtctact gctgtacat gatcatcttc accatggctg	2220
10	cctactacag gcccggtgat ggcttgcttc cctttaagat ggaaaaaact ggagactatt	2280
	tccgagttac tggagagatc ctgtctgtgt taggaggagt ctacttcttt ttccgaggga	2340
	ttcagtatct cctgcagagg cggccgtcga tgaagacct gtttggtggac agctacagtg	2400
	agatgctttt ctttctgcag tcaactgttca tgctggccac cgtggtgctg taacttcagcc	2460
15	acctcaagga gtatgtgggt tccatgggtat tctccctggc ctggggctgg accaactatgc	2520
	tctactacac ccgcggttcc cagcagatgg gcactctatgc cgtcatgata gagaagatga	2580
	tctctgagaga cctgtgcgtt ttcattgttg tctacatcgt cttcttgctc ggggtttcca	2640
	cagcgggtgg gacgtgatt gaagacggga agaatacttc cctgccgtct gagtccacgt	2700
	cgcacagggt gggggggcct gcctgcaggg cccccgatag ctccataaac agcctgtact	2760
20	ccacctgcct ggagctgttc aagttcacca tcggcatggg cgacctggag ttcactgaga	2820
	actatgactt caaggctgtc ttcactatcc tgctgctggc ctatgtaatt ctccactaca	2880
	tctctcgtct caacatgttc atcgccctca tgggtgagac tgtcaacaag atcgcacagg	2940
	agagcaagaa catctggaag ctgcagagag ccataccat cctggacacg gagaagagct	3000
25	tctttaagtg catgaggaag gccttccgct caggcaagct gctgcaggtg ggggtacacac	3060
	ctgatggcaa ggacgactac cgggtggtgt tcagggtgga cgaggtgaac tggaccacct	3120
	ggaacaccaa cgtgggcatc atcaacgaag acccgggcaa ctgtgagggc gtcaagcgca	3180
	ccctgagctt ctccctcggt tcaagcagag ttccaggcag acaactggaag aactttgccc	3240
30	tgggtccccc tttaagagag gcaagtgttc gagataggca gtctgctcag cccgagggaag	3300
	tttatctcgg acagttttca ggggtctctga agccagagga cgtgaggtc ttcaagagtc	3360
	ctgcegttcc cggggagaag tgaggacgtc acgcagacag cactgtcaac actgggcctt	3420
	aggagacccc gttgccacgg ggggtgctg agggaaacacc agtgcctctgt cagcagcctg	3480
35	gcctggtctg tgccctgcca gcatgttccc aaatctgtgc tggacaagct gtgggaagcg	3540
	ttcttggaag catggggagt gatgtacac caaccgtcac tgtccccaag tgaatctcct	3600
	aacagacttt cagggtttta ctcaactttac taaacagtkt ggatgggtcag tctctactgg	3660
	gacatgttag gcccttgctt tctttgattt tattcttttt tttagagacag aatttcactc	3720
40	ttctcaccca ggctggaatg cagtggcaca attttgctc cctgcaacct ccgcctctctg	3780
	gattccagca attctcctgc ctgggcttcc caagttagct ggattacagg cactgcccac	3840
	catgtctggc taattttttg tttttttt atagatatgg ggtttcgcca tgttggccag	3900
	gctggtctcg aactcctgac ctccaggtgat ccgcccacct cggcctccca aagtgtggg	3960
45	attacaggtg tgagcctcca caccctggctg tttctcttga ttttattctt tttctttttt	4020
	tctgtgagac agagtttcac tctgttgctc caggctggag tgcagtgggt tgatcttggt	4080
	tcaactgcaac ttctgcttcc cgggttcaag cgattcttct gcttcagtct cccaagtagc	4140
	ttggattaca ggtgagcact accacgcccg gctaattttt gtatttttaa taragacggg	4200
50	gtttcaccat gttggccagg ctgggtctcga actcttgacc tcaggtgac tgcccgcctt	4260
	ggcctcccaa agtgtggga ttacaggtgt gagccgtgc gctcggcctt ctttgatttt	4320
	atattatttag gagcaaaagt aaatgaagcc caggaaaaca cctttgggaa caaactcttc	4380

55

EP 0 943 683 A1

ctttgatgga aaatgcagag gcccttcttc tctgtgcegt gcttgctcct cttacctgcc 4440
 cgggtggttt gggggtgttg gtgttctctc cctggagaag atgggggagg ctgtccact 4500
 5 cccagctctg gcagaatcaa gctgttcgag cagtgccttc ttcctccttc cttacgatca 4560
 atcacagtct ccagaagatc agctcaattg ctgtgcaggt taaaactaca gaaccacatc 4620
 ccaaaggtag ctggttaagaa tgtttgaaag atcttccatt tctaggaacc ccagtcctgc 4680
 ttctccgcaa tggcacatgc ttccactcca tccatactgg cctcctcaaa taaacagata 4740
 10 tgtatacaat aaaaaaaaaa aaaaaaaaaa rrgcgccgcg tgaattctag acctgcccgg 4800
 gcg 4803

<210> 4
 <211> 839
 15 <212> PRT
 <213> Homo sapiens

<400> 4
 20 Met Lys Lys Trp Ser Ser Thr Asp Leu Gly Ala Ala Ala Asp Pro Leu
 1 5 10 15
 Gln Lys Asp Thr Cys Pro Asp Pro Leu Asp Gly Asp Pro Asn Ser Arg
 20 25 30
 25 Pro Pro Pro Ala Lys Pro Gln Leu Ser Thr Ala Lys Ser Arg Thr Arg
 35 40 45
 Leu Phe Gly Lys Gly Asp Ser Glu Glu Ala Phe Pro Val Asp Cys Pro
 50 55 60
 30 His Glu Glu Gly Glu Leu Asp Ser Cys Pro Thr Ile Thr Val Ser Pro
 65 70 75 80
 Val Ile Thr Ile Gln Arg Pro Gly Asp Gly Pro Thr Gly Ala Arg Leu
 85 90 95
 35 Leu Ser Gln Asp Ser Val Ala Ala Ser Thr Glu Lys Thr Leu Arg Leu
 100 105 110
 Tyr Asp Arg Arg Ser Ile Phe Glu Ala Val Ala Gln Asn Asn Cys Gln
 115 120 125
 40 Asp Leu Glu Ser Leu Leu Leu Phe Leu Gln Lys Ser Lys Lys His Xaa
 130 135 140
 Thr Asp Asn Glu Phe Lys Asp Pro Glu Thr Gly Lys Thr Cys Leu Leu
 145 150 155 160
 45 Lys Ala Met Leu Asn Leu His Asp Gly Gln Asn Thr Thr Ile Pro Leu
 165 170 175
 Leu Leu Glu Ile Ala Arg Gln Thr Asp Ser Leu Lys Glu Leu Val Asn
 180 185 190
 50 Ala Xaa Tyr Thr Asp Xaa Tyr Tyr Lys Gly Gln Thr Ala Leu His Ile
 195 200 205
 Ala Ile Glu Arg Arg Asn Met Ala Leu Val Thr Leu Leu Val Glu Asn

55

	210	215	220
	Gly Ala Asp Val Gln	Ala Ala His Gly Asp Phe Phe Lys Lys Thr	
5	225	230	235 240
	Lys Gly Arg Pro Gly Phe Tyr Phe Gly Glu Leu Pro Leu Ser Leu Ala		
	245	250	255
	Ala Cys Thr Asn Gln Leu Gly Ile Val Lys Phe Leu Leu Gln Asn Ser		
10	260	265	270
	Trp Gln Thr Ala Asp Ile Ser Ala Arg Asp Ser Val Gly Asn Thr Val		
	275	280	285
	Leu His Ala Leu Val Glu Val Ala Asp Asn Thr Ala Asp Asn Thr Lys		
15	290	295	300
	Phe Val Thr Ser Met Tyr Asn Glu Ile Leu Ile Leu Gly Ala Lys Leu		
	305	310	315 320
	His Pro Thr Leu Lys Leu Glu Glu Leu Thr Asn Lys Lys Gly Met Thr		
	325	330	335
20	Pro Leu Ala Leu Ala Ala Gly Thr Gly Lys Ile Gly Val Leu Ala Tyr		
	340	345	350
	Ile Leu Gln Arg Glu Ile Gln Glu Pro Glu Cys Arg His Leu Ser Arg		
	355	360	365
25	Lys Phe Thr Glu Trp Ala Tyr Gly Pro Val His Ser Ser Leu Tyr Asp		
	370	375	380
	Leu Ser Cys Ile Asp Thr Cys Glu Lys Asn Ser Val Leu Glu Val Ile		
	385	390	395 400
30	Ala Tyr Ser Ser Ser Glu Thr Pro Asn Arg His Asp Met Leu Leu Val		
	405	410	415
	Glu Pro Leu Asn Arg Leu Leu Gln Asp Lys Trp Asp Arg Phe Val Lys		
	420	425	430
35	Arg Ile Phe Tyr Phe Asn Phe Leu Val Tyr Cys Leu Tyr Met Ile Ile		
	435	440	445
	Phe Thr Met Ala Ala Tyr Tyr Arg Pro Val Asp Gly Leu Pro Pro Phe		
	450	455	460
40	Lys Met Glu Lys Thr Gly Asp Tyr Phe Arg Val Thr Gly Glu Ile Leu		
	465	470	475 480
	Ser Val Leu Gly Gly Val Tyr Phe Phe Phe Arg Gly Ile Gln Tyr Phe		
	485	490	495
45	Leu Gln Arg Arg Pro Ser Met Lys Thr Leu Phe Val Asp Ser Tyr Ser		
	500	505	510
	Glu Met Leu Phe Phe Leu Gln Ser Leu Phe Met Leu Ala Thr Val Val		
	515	520	525
50	Leu Tyr Phe Ser His Leu Lys Glu Tyr Val Ala Ser Met Val Phe Ser		
	530	535	540
	Leu Ala Leu Gly Trp Thr Asn Met Leu Tyr Tyr Thr Arg Gly Phe Gln		

55

545 550 555 560
 Gln Met Gly Ile Tyr Ala Val Met Ile Glu Lys Met Ile Leu Arg Asp
 5 565 570 575
 Leu Cys Arg Phe Met Phe Val Tyr Ile Val Phe Leu Phe Gly Phe Ser
 580 585 590
 Thr Ala Val Val Thr Leu Ile Glu Asp Gly Lys Asn Asp Ser Leu Pro
 10 595 600 605
 Ser Glu Ser Thr Ser His Arg Trp Arg Gly Pro Ala Cys Arg Pro Pro
 610 615 620
 Asp Ser Ser Tyr Asn Ser Leu Tyr Ser Thr Cys Leu Glu Leu Phe Lys
 15 625 630 635 640
 Phe Thr Ile Gly Met Gly Asp Leu Glu Phe Thr Glu Asn Tyr Asp Phe
 645 650 655
 Lys Ala Val Phe Ile Ile Leu Leu Leu Ala Tyr Val Ile Leu Thr Tyr
 20 660 665 670
 Ile Leu Leu Leu Asn Met Leu Ile Ala Leu Met Gly Glu Thr Val Asn
 675 680 685
 Lys Ile Ala Gln Glu Ser Lys Asn Ile Trp Lys Leu Gln Arg Ala Ile
 25 690 695 700
 Thr Ile Leu Asp Thr Glu Lys Ser Phe Leu Lys Cys Met Arg Lys Ala
 705 710 715 720
 Phe Arg Ser Gly Lys Leu Leu Gln Val Gly Tyr Thr Pro Asp Gly Lys
 725 730 735
 30 Asp Asp Tyr Arg Trp Cys Phe Arg Val Asp Glu Val Asn Trp Thr Thr
 740 745 750
 Trp Asn Thr Asn Val Gly Ile Ile Asn Glu Asp Pro Gly Asn Cys Glu
 755 760 765
 35 Gly Val Lys Arg Thr Leu Ser Phe Ser Leu Arg Ser Ser Arg Val Ser
 770 775 780
 Gly Arg His Trp Lys Asn Phe Ala Leu Val Pro Leu Leu Arg Glu Ala
 785 790 795 800
 40 Ser Ala Arg Asp Arg Gln Ser Ala Gln Pro Glu Glu Val Tyr Leu Arg
 805 810 815
 Gln Phe Ser Gly Ser Leu Lys Pro Glu Asp Ala Glu Val Phe Lys Ser
 820 825 830
 45 Pro Ala Ala Ser Gly Glu Lys
 835

<210> 5

50 <211> 432

<212> DNA

<213> Homo sapiens

55

<400> 5
 gagcttctcc ctgcggtcaa gcagagtttc aggcagacac tggaagaact ttgccctggt 60
 cccctttta agagaggcaa gtntcgana taggcagtct gtcagcccg aggaagtta 120
 tctgcgacag ttttcagggt ctctaaagcc agaggacgct gaggtcttca agagtcctgc 180
 cgcttcggg gagaagttag gacgtcacgc agacagcact gtcaacactg ggccttagga 240
 gaccccggtg ccacgggggg ctgctgaggg aacaccagtg ctttttcagc agccttgctt 300
 gggctctttgc ctgcccagca tgttcccaa tctgtgctgg acaagctgtg ggaagcggtt 360
 cttgggaagc atgggggagt gatgttacat ccaaccgtca ctgtcccaa gttgaatctt 420
 ccttaacaga tt 432

<210> 6
 <211> 65
 <212> PRT
 <213> Homo sapiens

<400> 6
 Ser Phe Ser Leu Arg Ser Ser Arg Val Ser Gly Arg His Trp Lys Asn
 1 5 10 15
 Phe Ala Leu Val Pro Leu Leu Arg Glu Ala Ser Xaa Arg Xaa Arg Gln
 20 25 30
 Ser Ala Gln Pro Glu Glu Val Tyr Leu Arg Gln Phe Ser Gly Ser Leu
 35 40 45
 Lys Pro Glu Asp Ala Glu Val Phe Lys Ser Pro Ala Ala Ser Gly Glu
 50 55 60
 Lys
 65

<210> 7
 <211> 3500
 <212> DNA
 <213> Homo sapiens

<400> 7
 cagcgtcggg tgcagtttgg ccggagggtg cagtgcagcag agattgcccc attgcactct 60
 agtctgggag acagggtgag acacacacac acagacacac acacacacac acacacacac 120
 acacaagcct aaacattcra ggccaggatg cttgacagat gttgattcat aaaaatgaca 180
 aaaagcacia aatccaaaat ctctataag ctgagtggtg gtggcagcga ggttgaagag 240
 caaaggcagg ccgggcacct ggctgatgat gtgtggacce gttgcacagc agggccccgc 300
 agtgcggtgt ggggtgtgggt gtgggtgggc cagtytctgc cgctcacctt attccaggga 360
 cacagtctgc ttggtcttct tggactgagc catctctatc accgagatcc tccctgaatt 420
 cagcccacga cagccacccc ggccgttttc cttgttctgt gtggggaggg aggcagcgcg 480

	gtgggttatca accacacct gcagaggagg cacctgaggg ccagagacga ggagggatgg	540
	gtctaaccaca gaaccacaga tggctctgag cggggggcct gtccaccctc ccaggccgac	600
5	gtcagtgggc gcaggactgc ctggggccctg ctaggccctgc tcacctctga ggctcttggg	660
	gtgagagggtt cagtcctgga aacacttcag ttctaggggg ctgggggcag cagcaagtctg	720
	gagttttggg gtacctctgt tcacaggggc cttggcaagg agggcagggtg gggcttaagg	780
	acaagcagtc cttactttgg gagtcaacct cggcgtgggtg gctgctgcag gttgcacact	840
10	gggccacaga ggatccagca aggatgaaga aatggagcag cacagacttg ggggcagctg	900
	cggaccacct ccaaaaggac acctgcccag accccctgga tggagacctt aactccaggc	960
	cacctccagc caagccccag ctctccacgg ccaagagccg caccggctc tttgggaagg	1020
	gtgactcggg ggaggcttcc cgggtggatt gccctcacga ggaaggtag ctggactcct	1080
15	ggccgacctt cacagtcagc cctgttatca ccatccagag gccaggagac gggcccaccg	1140
	gtgccaggct gctgtcccag gactctgtcg ccgccagcac cgagaagacc ctcaggctct	1200
	atgatcgag gagtatcttt gaagccgttg ctcagaataa ctgccaggat ctggagagcc	1260
	tgtctctctt cctgcagaag agcaagaagc acctcacaga caacgagttc aaagacctg	1320
20	agacagggaa gacctgtctg ctgaaagcca tgctcaacct gcacgacgga cagaacacca	1380
	ccatccccct gctcctggag atcgcgccgc aaacggacag cctgaaggag cttgtcaacg	1440
	ccagctacac ggacagctac tacaagggcc agacagcact gcacatcgcc atcgagagac	1500
	gcaacatggc cctgggtgacc ctctctggtg agaacggagc agacgtccag gctggggccc	1560
25	atggggactt ctttaagaaa accaaagggc ggccctggatt ctacttcggt gaactgcccc	1620
	tgtccctggc cgcgtgcacc aaccagctgg gcacgtgtaa gttcctgctg cagaactcct	1680
	ggcagacggc cgacatcagc gccagggact cgggtgggcaa cagggtgctg cagccctg	1740
	tggagggtgg cgacaacacg gccgacaaca cgaagtttgt gacgagcatg tacaatgaga	1800
30	ttctgatcct gggggccaaa ctgcaccoga cgtgaagct ggaggagctc accaacaaga	1860
	agggaaatgac ccgctgggtg ctggcagctg ggaccgggaa gatcggggtc ttggcctata	1920
	ttctccagcg ggagatccag gagcccgagt gcaggcaact gtccagggaag ttcaccgagt	1980
	gggcctacgg gcccgtgcac tcctcgctgt acgacctgtc ctgcatcgac acctgcgaga	2040
35	agaactcgtg gctggagggtg atcgccatca gcagcagcga gaccttaat cgccacgaca	2100
	tgtctctggt ggagccgctg aaccgactcc tgcaggacaa gtgggacaga ttcgtcaagc	2160
	gcactctcta cttcaacttc ctgggtctact gectgtacat gatcatcttc accatggctg	2220
	cctactacag ccccgtggat ggcttgccct cctttaagat ggaaaaaact ggagactatt	2280
40	tccgagttac tggagagatc ctgtctgtgt taggaggagt ctacttcttt ttccgaggga	2340
	ttcagtatct cctgcagagg cggccgtcga tgaagacctt gttgtggac agctacagtg	2400
	agatgctttt ctttctgcag tcaactgttca tgctggccac cgtgggtgctg tacttcagcc	2460
	acctcaagga gtatgtggct tccatgggtat tctcctggc cttgggctgg accaactg	2520
45	tctactacac ccgcgggttc cagcagatgg gcactctatgc cgtcatgata gagaagatga	2580
	tcctgagaga cctgtgccgt ttcatgtttg ttacgtcgt cttcttgttc gggttttcca	2640
	cagcgggtgt gacgtgatt gaagacggga agaatactc cctgccgtct gagtccacgt	2700
	cgcacagggt gcggggggcct gcctgcaggc ccccgatag ctccataaac agcctgtact	2760
50	ccacctgcct ggagctgttc aagttcacca tcggcatggg cgacctggag ttcactgaga	2820
	actatgactt caaggctgtc ttcatcatcc tgcgtctggc ctatgtaatt ctcacctaca	2880
	tcctctgctt caacatgctc atcgccctca tgggtgagac tgtcaacaag atcgcacagg	2940
	agagcaagaa catctggaag ctgcagagag ccatcaccat cctggacacg gagaagagct	3000

55

5 tccttaagtg catgaggaag gccttccgct caggcaagct gctgcagggtg gggtagacac 3060
 ctgatggcaa ggacgactac cggtaggtgct tcaggggtgga cgaggtgaac tggaccacct 3120
 ggaacaccaa cgtgggcatc atcaacgaag acccgggcaa ctgtgagggc gtcaagcgca 3180
 ccctgagctt ctccctgcgg tcaagcagag ttccaggcag aacttggaag aactttgccc 3240
 tggtcccccct ttttaagagag gcaagtgtc gagataggca gtctgtctcag cccgaggaag 3300
 tttatctgcy acagttttca gggctctctga agccagagga cgtgaggtc ttcaagagtc 3360
 10 ctgccgcttc cggggagaag tgaggacgtc acgcagacag cactgtcaac actgggcctt 3420
 aggagacccc gttgccacgg ggggtctgtg agggaaacac agtgetctgt cagcagcctg 3480
 gcctggtctg tgctgcccc 3500

15 <210> 8
 <211> 839
 <212> PRT
 <213> Homo sapiens

20 <400> 8
 Met Lys Lys Trp Ser Ser Thr Asp Leu Gly Ala Ala Ala Asp Pro Leu
 1 5 10 15
 25 Gln Lys Asp Thr Cys Pro Asp Pro Leu Asp Gly Asp Pro Asn Ser Arg
 20 25 30
 Pro Pro Pro Ala Lys Pro Gln Leu Ser Thr Ala Lys Ser Arg Thr Arg
 35 40 45
 30 Leu Phe Gly Lys Gly Asp Ser Glu Glu Ala Phe Pro Val Asp Cys Pro
 50 55 60
 His Glu Glu Gly Glu Leu Asp Ser Cys Pro Thr Ile Thr Val Ser Pro
 65 70 75 80
 35 Val Ile Thr Ile Gln Arg Pro Gly Asp Gly Pro Thr Gly Ala Arg Leu
 85 90 95
 Leu Ser Gln Asp Ser Val Ala Ala Ser Thr Glu Lys Thr Leu Arg Leu
 100 105 110
 40 Tyr Asp Arg Arg Ser Ile Phe Glu Ala Val Ala Gln Asn Asn Cys Gln
 115 120 125
 Asp Leu Glu Ser Leu Leu Leu Phe Leu Gln Lys Ser Lys Lys His Leu
 130 135 140
 45 Thr Asp Asn Glu Phe Lys Asp Pro Glu Thr Gly Lys Thr Cys Leu Leu
 145 150 155 160
 Lys Ala Met Leu Asn Leu His Asp Gly Gln Asn Thr Thr Ile Pro Leu
 165 170 175
 50 Leu Leu Glu Ile Ala Arg Gln Thr Asp Ser Leu Lys Glu Leu Val Asn
 180 185 190
 Ala Ser Tyr Thr Asp Ser Tyr Tyr Lys Gly Gln Thr Ala Leu His Ile
 195 200 205

55

Ala Ile Glu Arg Arg Asn Met Ala Leu Val Thr Leu Leu Val Glu Asn
 210 215 220
 Gly Ala Asp Val Gln Ala Ala Ala His Gly Asp Phe Phe Lys Lys Thr
 225 230 235 240
 Lys Gly Arg Pro Gly Phe Tyr Phe Gly Glu Leu Pro Leu Ser Leu Ala
 245 250 255
 Ala Cys Thr Asn Gln Leu Gly Ile Val Lys Phe Leu Leu Gln Asn Ser
 260 265 270
 Trp Gln Thr Ala Asp Ile Ser Ala Arg Asp Ser Val Gly Asn Thr Val
 275 280 285
 Leu His Ala Leu Val Glu Val Ala Asp Asn Thr Ala Asp Asn Thr Lys
 290 295 300
 Phe Val Thr Ser Met Tyr Asn Glu Ile Leu Ile Leu Gly Ala Lys Leu
 305 310 315 320
 His Pro Thr Leu Lys Leu Glu Glu Leu Thr Asn Lys Lys Gly Met Thr
 325 330 335
 Pro Leu Ala Leu Ala Ala Gly Thr Gly Lys Ile Gly Val Leu Ala Tyr
 340 345 350
 Ile Leu Gln Arg Glu Ile Gln Glu Pro Glu Cys Arg His Leu Ser Arg
 355 360 365
 Lys Phe Thr Glu Trp Ala Tyr Gly Pro Val His Ser Ser Leu Tyr Asp
 370 375 380
 Leu Ser Cys Ile Asp Thr Cys Glu Lys Asn Ser Val Leu Glu Val Ile
 385 390 395 400
 Ala Tyr Ser Ser Ser Glu Thr Pro Asn Arg His Asp Met Leu Leu Val
 405 410 415
 Glu Pro Leu Asn Arg Leu Leu Gln Asp Lys Trp Asp Arg Phe Val Lys
 420 425 430
 Arg Ile Phe Tyr Phe Asn Phe Leu Val Tyr Cys Leu Tyr Met Ile Ile
 435 440 445
 Phe Thr Met Ala Ala Tyr Tyr Arg Pro Val Asp Gly Leu Pro Pro Phe
 450 455 460
 Lys Met Glu Lys Thr Gly Asp Tyr Phe Arg Val Thr Gly Glu Ile Leu
 465 470 475 480
 Ser Val Leu Gly Gly Val Tyr Phe Phe Phe Arg Gly Ile Gln Tyr Phe
 485 490 495
 Leu Gln Arg Arg Pro Ser Met Lys Thr Leu Phe Val Asp Ser Tyr Ser
 500 505 510
 Glu Met Leu Phe Phe Leu Gln Ser Leu Phe Met Leu Ala Thr Val Val
 515 520 525
 Leu Tyr Phe Ser His Leu Lys Glu Tyr Val Ala Ser Met Val Phe Ser
 530 535 540

EP 0 943 683 A1

	Leu	Ala	Leu	Gly	Trp	Thr	Asn	Met	Leu	Tyr	Tyr	Thr	Arg	Gly	Phe	Gln
	545					550					555					560
5	Gln	Met	Gly	Ile	Tyr	Ala	Val	Met	Ile	Glu	Lys	Met	Ile	Leu	Arg	Asp
						565					570					575
	Leu	Cys	Arg	Phe	Met	Phe	Val	Tyr	Val	Val	Phe	Leu	Phe	Gly	Phe	Ser
						580					585					590
10	Thr	Ala	Val	Val	Thr	Leu	Ile	Glu	Asp	Gly	Lys	Asn	Asp	Ser	Leu	Pro
						595					600					605
	Ser	Glu	Ser	Thr	Ser	His	Arg	Trp	Arg	Gly	Pro	Ala	Cys	Arg	Pro	Pro
						610					615					620
15	Asp	Ser	Ser	Tyr	Asn	Ser	Leu	Tyr	Ser	Thr	Cys	Leu	Glu	Leu	Phe	Lys
						625					630					635
	Phe	Thr	Ile	Gly	Met	Gly	Asn	Leu	Glu	Phe	Thr	Glu	Asn	Tyr	Asn	Phe